Undertaking #27

MPI to produce the final report regarding the analysis of data collected from the Ready Assess Pilot Project.

RESPONSE:

Please see *Appendix 1 – Driver Testing and Assessment Pilot Program Oct 31, 2022*.

Final Report for the Driver Testing and Assessment Pilot (DTAP) program: Testing an established virtual driving test platform in Manitoba Public Insurance Corporation's road testing workflow

Submitted to: Manitoba Public Insurance Corporation
Prepared by: Diagnostic Driving, Inc.
Finalized: October 21, 2022

TABLE OF CONTENTS

Suggested Citation:

Kandadai, V, Sosnowski, M. (2022). Final Report for the Driver Testing and Assessment Pilot (DTAP) program: Testing an established virtual driving test platform in Manitoba Public Insurance Corporation's road testing workflow. Diagnostic Driving, Inc. Philadelphia, PA, USA.

Author Information

	Venk Kandadai, MPH Co-founder and CEO Diagnostic Driving, Inc.
	Michael Sosnowski Director, Client Success Diagnostic Driving, Inc.

Disclosures and Signatures

This report details the results of a 11-month pilot program in which a novel virtual driving test program was integrated into Manitoba Public Insurance Corporation's (MPI) driver testing workflow with customers seeking independent driving privileges. The results contained in this report are based on a good-faith analysis of the data collected during this pilot, using standard best practices for data collection, data management, and data analysis. Diagnostic Driving, Inc. (DDI) is the developer of the virtual driving test (commercially referred to as Ready-Assess™). The author of this report, Venk Kandadai's potential of conflict of interest is managed by a conflict management plan from DDI whereby Mr. Kandadai and DDI staff had no direct interactions with MPI customers. All field data collection procedures were carried out by MPI staff, and all methods and analyses were reviewed and approved by an outside consultant with no intellectual or financial interest in DDI (Nicolas Skuli, PhD; a Senior Research Investigator and Director of the Stem Cell and Xenograft Core at the University of Pennsylvania).

Official report prepared by:

Name: Venk Kandadai, MPH

Title: Co-founder and CEO

Organization: Diagnostic Driving, Inc.

Address: 705 S. 50th Street, Philadelphia, PA, 19143, USA

Contact Email: info@diagnosticdriving.com

Date: October 21, 2022

Reviewed and accepted as the final contractual deliverable by:

Name: Patrick Sarginson

Title: Director, DVA Policy and Registrar of Motor Vehicles

Organization: Manitoba Public Insurance Corporation

Address: 702–234 Donald Street Box 6300, Winnipeg, Manitoba, R3C 4A4, Canada

Contact Email: psarginson@mpi.mb.ca

Date: October 24, 2022

Patrick Sarginson

EXECUTIVE SUMMARY

Particularly for young drivers, the point of licensure provides a critical safety-intervention opportunity: crash risk peaks immediately after licensure and declines during the following two years. Additionally, underprepared applicants of all ages attempting a road skills examination (RSE) for licensure (1) create operational burdens for motor vehicle administrations; and (2) pose public safety risks to themselves, license examiners and to the broader public.

Between 2021–2022, the Manitoba Public Insurance Corporation (MPI) commissioned an 11-month pilot to test a well-established virtual driving test (VDT) platform in their road-testing workflow with customers scheduled to attempt a RSE. The pilot's **success criteria** were to:

1. Demonstrate initial feasibility of integrating the VDT into MPI's driver testing workflow with customers
2. Demonstrate a low frequency of simulator-based motion sickness (SMS) symptoms after completing the VDT
3. Establish initial evidence that performance in the VDT is associated with RSE outcomes:
    a. <u>Note</u>: MPI chose not to develop a customized VDT scoring algorithm optimized for customer screening prior to the attempting the RSE

Immediately before their scheduled RSE, 967 customers voluntarily attempted the self-guided VDT at five MPI Service Centre locations. The results from the pilot established foundational evidence across all the success criteria. Specifically:

- **Performance in the VDT was predictive and generalizable:** Performance in the VDT was predictive of the customer's RSE result, consistent with previous findings in other motor vehicle administrations outside of Manitoba (Ohio and Delaware).
    - There was a compounding effect of the VDT Error Score on the likelihood of failing the RSE - as the number of significant errors in the VDT increased, the likelihood of failing the RSE significantly increased (*Any RSE attempt*: aOR=1.39, 95% CI: 1.20-1.61, $p<0.0001$; *First RSE attempt*: aOR=1.64, 95% CI: 1.28-2.13, $p=0.0001$).
- **SMS symptoms were infrequent and mostly minor.**
- **Customers found the VDT realistic and relevant – providing an opportunity to deliver high quality feedback and interventions targeting skill gaps for all customers seeking licensure.**
- The VDT was **feasible to integrate** into a busy road-testing workflow.
    - As expected, most customers spent less than 15 minutes independently completing the VDT. No schedule backlogs were reported.
    - Most customers completed the VDT in a self-directed manner without staff supervision or intervention.
    - Reported support requests were few and minor (two support requests over 11 months → both resolved in less than 1 hour).

The VDT can bolster the safety-intervention opportunity provided when MPI's customers seek licensure. Furthermore, the results from this pilot support the immediate readiness of the VDT and the benefits of integrating it into a safety-screening paradigm for the RSE (and point of licensure). The implications of the pilot's results and recommendations for a larger-scale deployment are further discussed in this report.

SIGNIFICANCE, CONTEXT AND PILOT SUCCESS CRITERIA

Significance and context:

In Canada, motor vehicle crashes continue to be a leading cause of death and injury, particularly among adolescents and young adults (Transport Canada, National Collision Database, 2020). Furthermore, newly licensed young drivers are at a higher risk for being involved in motor vehicle crashes, with average crash rates peaking in the months immediately after licensure and then declining to rates comparable to adult-experienced drivers over the first few years of licensed driving (Curry et al., 2015, 2017; Tefft, 2017). Additionally, underprepared drivers attempting a road skills examination (RSE) to seek licensure pose public safety risks to themselves, license examiners and to the broader public.

In 2020, Manitoba Public Insurance Corporation (MPI) contacted Diagnostic Driving, Inc. (DDI) regarding their virtual driving test (VDT) product, Ready-Assess™. At the time, MPI was considering a solution to (1) help mitigate issues around customers using multiple attempts to pass the road skills examination (RSE); and (2) manage the safety concerns of having high fail rates on the RSE due to underprepared customers - a public safety concern to MPI's examiners, customers, and the broader community in Manitoba.

In 2021:

- Success criteria were defined and mutually agreed to by both parties (MPI and DDI).
- A master service agreement was executed between DDI and MPI to conduct a pilot addressing the defined success criteria.
- Data collection commenced in September with RSE customers, lasting through July 2022.

About the virtual driving test:

The VDT (commercially referred to as Ready-Assess™) is a well-established and evidence-based virtual driving test platform used in both commercial and clinical settings (Walshe et al., 2022, Grethlein et al., 2022, Grethlein et al., 2020, Lee, 2021, Walshe et al., 2020; Winston et al., 2019). <u>The objective of the VDT is to stress test drivers in ecologically valid crash scenarios</u> (McDonald et al., 2014), known to cause severe injury to both adolescent and adult drivers. The crash scenarios in the VDT are often difficult to replicate during a formal RSE with an examiner or during a typical behind-the-wheel driver instruction program (due to safety concerns and the inconsistency of safely exposing drivers to these conditions). At the end of the VDT, drivers receive automated

Figure 1: A typical VDT workstation.

Final Report: Page 5

and personalized feedback on their performance, incorporating how they did compared to their peers, 'teachable-moment' guidance leveraging observed unsafe driving behaviors during the VDT, and a personalized remediation plan to address areas for improvement - utilizing an evidence-based intervention proven to improve safe driving performance (Mirman et al, 2014). Designed to be implemented in limited-resource settings, the VDT:

- Uses a self-directed workflow lasting approximately 15 minutes.
- Runs on inexpensive and off-the-shelf equipment, scalable for broad reach.
- Requires minimal space (e.g., a small office desk or cubicle) and minimal Internet connectivity speeds (wired or wireless).

Of note, driving performance in the VDT was demonstrated to be predictive of RSE outcomes (for ascertaining independent driving privileges) in jurisdictions in Ohio (Ohio Bureau of Motor Vehicles, Ohio, USA) and in Delaware (Delaware Division of Motor Vehicles, Delaware, USA).

DTAP program success criteria:

MPI developed criteria to gauge the success of the DTAP program and to inform their decisions regarding larger-scale implementation of the VDT towards relevant applications. These criteria included:

1. Demonstrating initial feasibility of integrating the VDT into MPI's driver testing workflow with customers
2. Demonstrating that the frequency of simulator-based motion sickness symptoms after completing the VDT is low
3. Establishing initial evidence that performance in the VDT is associated with RSE outcomes

METHODS

Overview:

From September through December 2021, MPI installed VDT workstations across four service centre locations (Bison, Brandon, Gateway, and St. Mary's). A fifth workstation was installed at the Main Street service centre in April 2022. MPI staff were trained in August 2021 by DDI team members on the VDT standard operating procedures, using a train-the-trainer methodology. Data collection with actual MPI customers commenced in September 2021 and ended on July 31, 2022.

Pilot design:

In a cross-sectional design, MPI customers scheduled for their road skills examination (RSE), were asked to independently complete the VDT immediately before their scheduled RSE (on a voluntary, non-punitive basis) on a workstation installed at the service centre. Regardless of VDT performance, all customers were invited to complete their scheduled RSE. To address the success criteria defined at the start of the pilot program, both the customers and the RSE examiners were blinded to VDT results (as the VDT is capable of automatically outputting feedback containing results). This was done to mitigate testing bias and anxiety associated with completing the RSE. Customers commenced a VDT session by inputting their unique Service Number (SN) into the VDT software installed on the workstation. The SN was a unique identifier to the customer and was subsequently used to link a customer's VDT results with their RSE result, RSE attempt history, and demographic information (age and gender).

Virtual driving test procedures for customers:

Figure 2: The VDT installed in a typical driver testing center (photos courtesy of the Ohio Bureau of Motor Vehicles).

Upon logging into the VDT software installed on the workstation, each customer was exposed to three primary modules in a self-directed manner:

1. In the first module, an animated orientation video was presented, lasting approximately 30 seconds. The purpose of this video was to provide the customer with a basic introduction to the VDT.
2. The second module incorporated an introductory drive. During this module, the customer was introduced to the different VDT control inputs (e.g., steering, brake, throttle, turn signal, transmission, etc.). Additionally, the customer was asked to complete a series of basic driving maneuvers (accelerating, braking, steering through a curve and executing a 90-degree turn) to acclimate themselves to the program. This module lasted approximately 3 minutes and allowed the customer to repeat the process, if needed.
3. The third module was a virtual driving test route which was scored for performance. The route contained variations of common crash scenarios as described in McDonald et al. (2014) and common driving tasks and roadway hazards (e.g., pedestrian crosswalks, school zones, ambulances, etc.), embedded in a typical driving route, and lasting approximately 10 minutes. Each customer was exposed to the same VDT driving route (same tasks and exposures in the same order).

A fourth module (customer debriefing) was appended to the VDT workflow specifically for the purposes of this pilot. In this final module, the customer was asked a series of on-screen questions regarding the relevance of the VDT and reporting self-reported symptoms of simulator-based motion sickness (if any). The speedometer used during the VDT displayed in kilometers-per-hour and all customers had the option to receive voiceover prompts in either English or French at the start of the VDT session. Upon completing the VDT, data was automatically uploaded to DDI's cloud service for storing and processing.

Data preparation:

VDT data processing:
Each completed VDT session generated a JSON file (referred to as a replay file) containing a 10 Hz multivariate time series record. The replay file was used to compute unique time series channels to measure driving performance across multiple driving elements (e.g., speed, lane deviation, following distance, elapsed time, etc.). Within the time series channels, eye-tracking channels were also computed (using an automated, time-synchronized integration procedure developed by Diagnostic Driving, Inc.) to quantify eye-gaze movement during the VDT session. These channels were ultimately converted into unique variables (significant driving errors and infractions). Additionally, a composite score of VDT performance was automatically computed (VDT Error Score) which was a weighted, linear combination of significant driving errors, made along the driving route (e.g., colliding with vehicles, colliding with pedestrians, red light errors, etc.). The VDT Error Score took on integer values (0 → infinity), where a larger score indicated poorer driving performance. The VDT Error Score estimates overall driving performance during the VDT, and it was previously used to predict RSE results for licensure (Winston et al., 2019). Additionally, it serves as a practical and convenient way to summarize performance, be easily interpreted, and be displayed back to the driver in the form of feedback containing practical guidance.

RSE data provided by MPI and used in the analysis:
MPI sent a final Excel file to DDI containing the following data, indexed by SN:
- Customer’s integer age (in years) on the date of the customer’s VDT session
- Customer’s gender
- Customer’s most recent RSE result (Pass/Fail), if available
- Result effective date (date when most recent RSE was attempted)
- # of previous RSE attempts made by the customer prior to the most recent RSE

DDI was not involved in generating the RSE data provided by MPI. All RSE data used for analysis purposes was used in ‘as-is’ condition by DDI without any additional modifications. It was assumed that the ‘*result effective date’* was the same date that the customer attempted the VDT. There was no personally identifiable information about customers transferred between MPI and DDI. Additionally, DDI team members (staff, consultants, or representatives) had no direct interactions with MPI customers.

Additional data collected:
The debriefing module at the end of the VDT session incorporated four additional questions that the customer answered. Customers would only respond to these questions if they completed the entire VDT workflow (as the questions were displayed at the end of the session). The questions covered two broad categories detailed in Table 1.

Table 1: Debriefing questions asked to customers upon completing the VDT.

Item	Ascertaining	Response Choices
The driving situations I just encountered are a reasonable representation of what I see on the road	Perceived relevance	5-pt Likert scale (1=Strongly disagree; 5=Strongly agree)
I am curious to see how I did in this simulated driving test.	Perceived relevance	5-pt Likert scale (1=Strongly disagree; 5=Strongly agree)
How nauseous are you feeling right now?	Symptoms simulator-based motion sickness	Not at all; A little bit but it’s not bothering me; More than normal; Extremely
How dizzy are you feeling right now?	Symptoms simulator-based motion sickness	Not at all; A little bit but it’s not bothering me; More than normal; Extremely

Final analytical dataset:

Data from the RSE data transfer, the VDT, and the customer debriefing, were linked together (indexed by SN) to produce a final dataset used for the analysis. The primary outcome variable was *RSE result* (Pass/Fail) and the primary predictor variable was the *VDT Error Score* (log transformed to meet normality assumptions for predictive modeling).

Analytical methods utilized:

Descriptive summary statistics (means, medians, ranges, proportions, correlations) were used where appropriate to summarize data pertaining to feasibility, customer relevance, simulator-based motion sickness, and VDT errors and infractions. Logistic regression was performed to test the efficacy hypothesis that performance in the VDT is associated with RSE outcomes. Model adjustments were made to control for the customer’s age, gender, and the service centre location where they completed the VDT and RSE. Adjusted odds ratios, 95% confidence intervals, and p-values were reported. A threshold of $p<0.05$ was used to ascertain statistical significance and all analyses were performed using R-software (version 4.0.2; r-project.org).

RESULTS

Sample overview:

Between September 2021 and July 2022, 967 customers attempted the VDT. Among the 967 customers who attempted the VDT, 823 customers (85%) completed the VDT workflow from end-to-end (fully completed the virtual driving test route). Table 2 below depicts the VDT sessions administered by service centre location. Of note, the overall VDT completion rate (85%) started lower than anticipated earlier in the pilot when the program was new to both MPI staff and customers. As the pilot progressed, the VDT completion rate significantly increased towards the later months of the pilot (>90%). Additionally, we found moderate evidence that not completing the VDT was associated with failing the RSE (Odds Ratio: 1.55; 95% CI: 1.05-2.30; p=0.03) - as previously observed by Winston et al., 2019. DDI and MPI did not document the specific reasons for why the 144 customers did not fully complete the VDT workflow - as the customers attempted the VDT independently without supervision.

Table 2: Customer VDT sessions across MPI service centre locations.

	Service Centre Location					
	All locations	Bison	Brandon	Gateway	Main St.	St. Mary's
Total VDT sessions	967	395	45	181	88	258
Completed VDT sessions	823	300	40	162	79	242
Completion rate	85%	76%	89%	90%	90%	94%

Among the 967 customers, 430 were female (45%), 527 were male (54%), and 10 (1%) were unknown. Nine customers did not have an age value provided by MPI. Among the 958 customers whose age was known at the time of attempting the VDT, the mean age was 26.6 years (minimum=16; median=23; sd=10.4; maximum=74). Additionally, the RSE fail rate of this sample was 63% (regardless of the number of previous RSE attempts), 43% (for customers on their first attempt at the RSE), and 75% (for customers who had previously attempted the RSE at least once). MPI subject matter experts indicated that the sample collected during the piloting period was representative of the overall population that MPI sees for RSE services.

Success Criteria 1 *- Demonstrate feasibility of integrating the VDT into MPI's driver testing workflow with customers*:

Reliability of service:
The 967 customer VDT sessions represented 221 cumulative hours of VDT operating time. During this period, two support requests were submitted to DDI's Help Desk Support service and categorized by their severity level, as determined by DDI's Service Level Agreements with MPI (refer to Table 3).

Table 3: Support requests submitted by MPI to DDI's Help Desk during the piloting period – over 221 hours of cumulative VDT operating time.

Severity Level	Definition	# Requests Submitted	Incident Rate (per 1000 VDT sessions)
Level 1	The DDI software is not available during operating hours	0	0
Level 2	The DDI software is available, but functionality is degraded or restricted, adversely impacting business during normal operating hours	0	0
Level 3	Any issue with the DDI software or with the DDI Web Application software that does not prevent business services from operating the DDI software within 24 hours of identifying the event	2 (all resolved in <1 hour)	2.1

Note: One incident was not classified as a support request as it was determined that MPI's firewall was blocking DDI's cloud services at pilot launch. This was not due to any issue related to the DDI software and was resolved internally by MPI before data collection began with customers.

After completing the 1-hour virtual staff training session with DDI, which incorporated an in-depth walkthrough of the VDT standard operating procedures, no other issues were reported beyond those identified above. Also, there were no schedule backlogs reported during the piloting period. Finally, there was no loss of data among the 823 completed VDT sessions when automatically uploaded to DDI's cloud servers for processing, upon completing a VDT session.

Customer time spent during a VDT session:
The average VDT session for a customer lasted 13.7 minutes, which was within range of the planned 15-minute benchmark. Additionally, 75% of all customers spent less than 15 minutes in a VDT session. Table 4 depicts the distribution of customer session time in the VDT.

Table 4: Distribution of customer time spent during VDT sessions (in minutes).

Minimum	Q1	Median	Mean (sd)	Q3	Maximum
0.75	12.2	13.4	13.7 (3.3)	14.9	33.5

Relevance to MPI customers:
Although customers did not receive a feedback report upon completing the VDT (a requirement for the pilot design), they were asked to respond to a series of questions regarding the relevance of the VDT. Among the 823 customers that completed a VDT session, 803 (98%

response rate) responded to the questions pertaining to the relevance of the VDT. Table 5 depicts these results.

Table 5: Customers' self-reported responses pertaining to the perceived relevance of the VDT.

Customer Response: N (%)	Numeric value	The driving situations I just encountered are a reasonable representation of what I see on the road.	I am curious to see how I did in this simulated driving test.
Strongly agree	5	131 (16%)	155 (19%)
Agree	4	494 (62%)	342 (43%)
No opinion	3	57 (7%)	194 (24%)
Disagree	2	70 (9%)	58 (7%)
Strongly disagree	1	51 (6%)	54 (7%)
Numeric Median Response		4.0	4.0
Numeric Mean Response		3.7	3.6

Most customers (78%) 'agreed' or 'strongly agreed' that the driving situations in the VDT were a reasonable representation of what they see on the road. Additionally, most customers (62%) were receptive to seeing how they performed in the VDT.

***Success Criteria 2** - Demonstrate that the frequency of simulator-based motion sickness symptoms after completing the VDT is low*:

Among the 823 customers that completed a VDT session, 803 (98% response rate) responded to the questions pertaining to symptoms of motion sickness. Table 6 depicts the customers' self-reported responses to each question.

Table 6: Customers' self-reported symptoms of simulator-based motion sickness immediately after completing the VDT.

Customer Response: N (%)	Numeric Value	How **nauseous** are you feeling right now?	How **dizzy** are you feeling right now?
Not at all	1	421 (52%)	546 (68%)
A little but it's not bothering me	2	265 (33%)	178 (22%)
More than normal	3	99 (12%)	70 (9%)

Extremely	4	18 (2%)	9 (1%)
Numeric Median Response		1.0	1.0
Numeric Mean Response		1.6	1.4

For each symptom (nausea and dizziness), 85% and 90% of customers reported ‘Not at all’ or ‘A little but it’s not bothering me’, respectively. These results were generally consistent with data captured from jurisdictions outside of Manitoba.

Table 7: Cross-tabulation of self-reported symptoms of nausea and dizziness after completing the VDT.

	Dizzy			
Nauseous	Not at all	A little but it’s not bothering me	More than normal	Extremely
Not at all	387 (48%)	31	3	0
A little but it’s not bothering me	129	114 (14%)	21	1
More than normal	25	30	40 (5%)	4
Extremely	5	3	6	4 (0.5%)

The customers’ self-reported responses to each symptom (nauseous and dizzy) were statistically correlated with each other ($r=0.6$; $p<0.0001$). A cross-tabulation (as shown in Table 7) revealed that (1) nearly half of the customers did not experience any symptoms; and (2) less than 1% of customers experienced ‘extreme’ levels of both symptoms.

Success Criteria 3 - *Establish initial efficacy that performance in the VDT is associated with RSE outcomes*:

<u>Derivation of analytic sample</u>:
Of the 967 customer VDT sessions, 931 had a RSE result provided by MPI that was recorded within two days of attempting the VDT. Furthermore, an additional 140 customer VDT sessions were removed from the analysis because the customer did not complete the VDT (unable to generate a VDT Error Score); resulting in a final sample size of N=791. Figure 3 depicts the derivation of the sample.

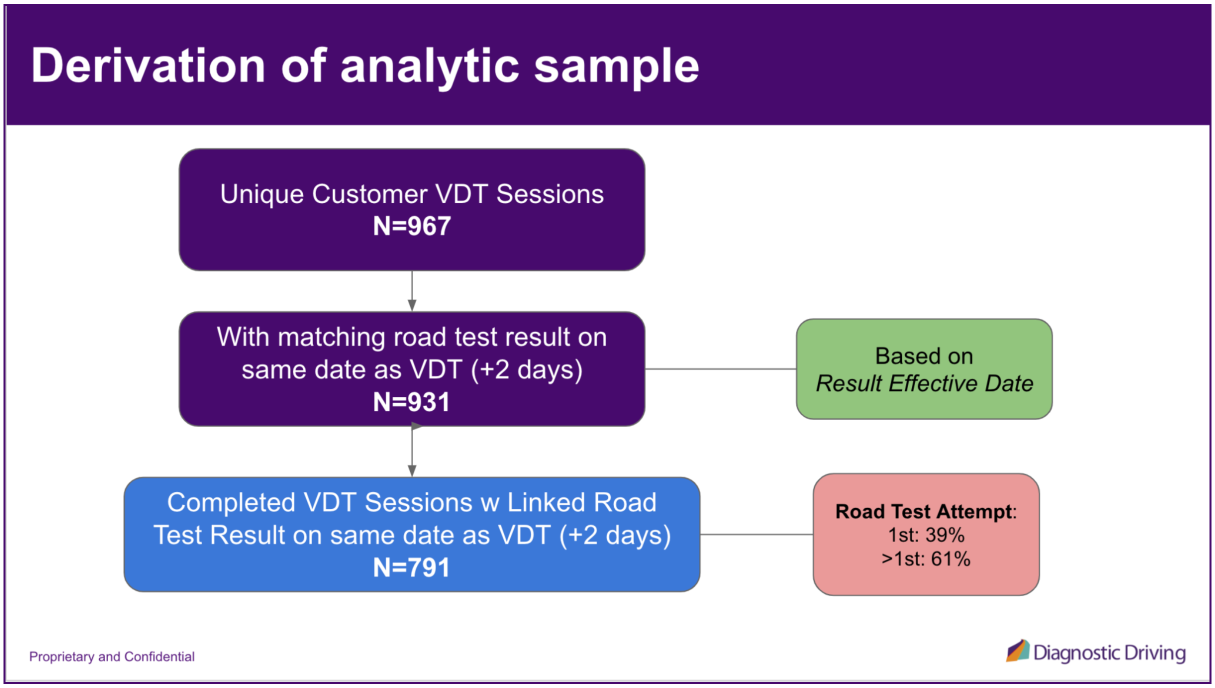

Figure 3: Derivation of the analytic sample used to conduct the efficacy analysis (Success criteria #3).

<u>Efficacy result - VDT Error Score and the likelihood of failing the RSE</u>:
We fit a logistic regression model to predict the RSE outcome using the log-transformed VDT Error Score, while controlling for the customer's age, gender, and the service centre location where they completed the RSE. The log-transformed VDT Error Score was used to meet normality assumptions to fit a logistic regression model, where:

- Log-transformed VDT Error Score = ln(VDT Error Score+1)
- VDT Error Score = $e^{(\text{log-transformed VDT Error Score})}-1$

Table 8 depicts the results of the logistic regression model for (1) all customers regardless of the amount of previous RSE attempts they had; and (2) only those customers who were on their first RSE attempt.

Table 8: Logistic regression results for log-transformed VDT Error Score predicting RSE results. Adjusted odds ratios were reported and controlled for the customer's age, gender, and the Service Centre Location where they completed the VDT and the RSE. For example, an adjusted odds ratio of 1.39 implies a 39% increase in the likelihood of failing the RSE for every 1-unit increase in log-transformed VDT Error Score.

Condition	Reference Outcome	Adjusted Odds Ratio (95% CI)	P-value
Any RSE attempt	RSE=Fail	1.39 (1.20-1.61)	<0.0001
First RSE attempt	RSE=Fail	1.64 (1.28-2.13)	0.0001

VDT Error Score (log-transformed) was significantly associated with the RSE outcome regardless of if it was the customer's first attempt at the RSE or not. Specifically, there was a compounding effect of the VDT Error Score on the likelihood of failing the RSE – as the number of significant errors in the VDT increased, the likelihood of failing the RSE increased (refer to Figure 4 below).

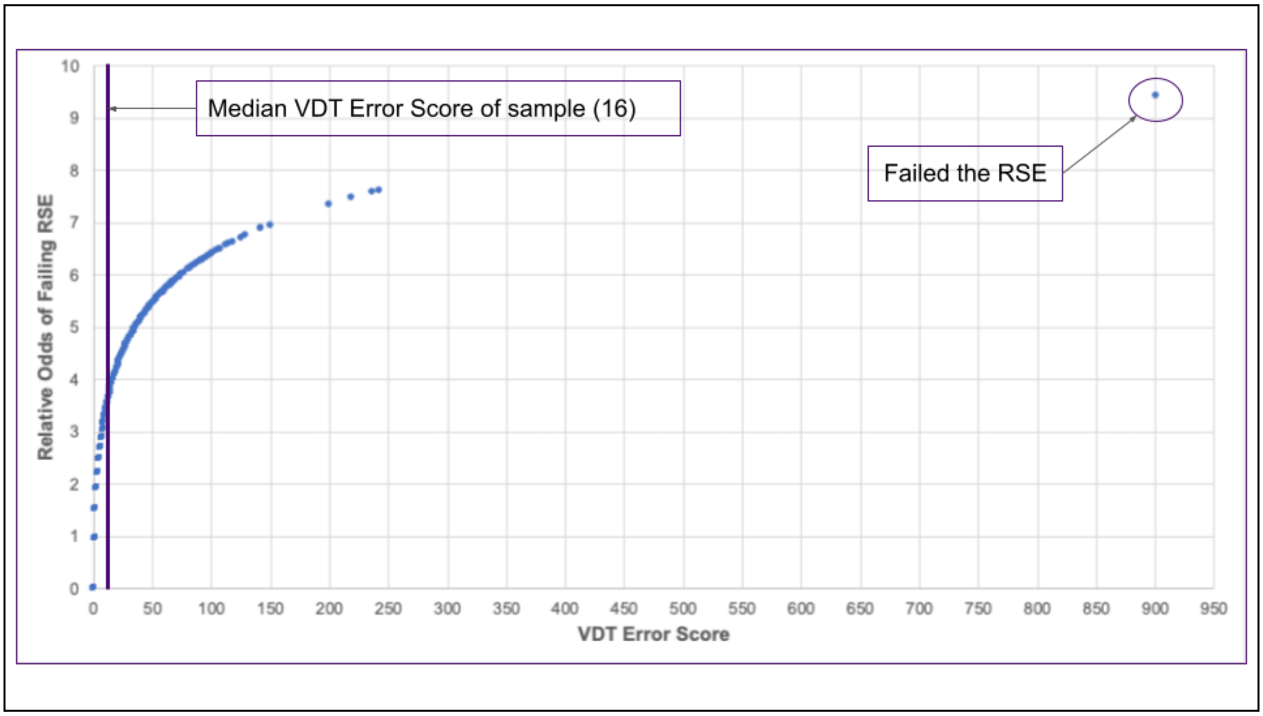

Figure 4: VDT Error Score and the relative odds of failing the RSE. For example, a VDT Error Score of 150 increased the likelihood of failing the RSE by a factor of ~7x. Note: Log-transformed VDT Error Score was transformed back to VDT Error Score for interpretation and graphical purposes. Additionally, one customer (circled) had a VDT Error Score>900 which was an extreme outlier of poorer performance (this customer also failed the RSE). There were negligible differences in the computed Odds Ratios, 95% CIs, and P-values when examining the results after removing this customer.

We also examined VDT Error Score and RSE "Pass" rates within VDT Error Score groupings (refer to Figure 5). Poorer VDT Error Scores were associated with lower pass rates on the RSE. This result was consistent regardless of the RSE attempt (first or any).

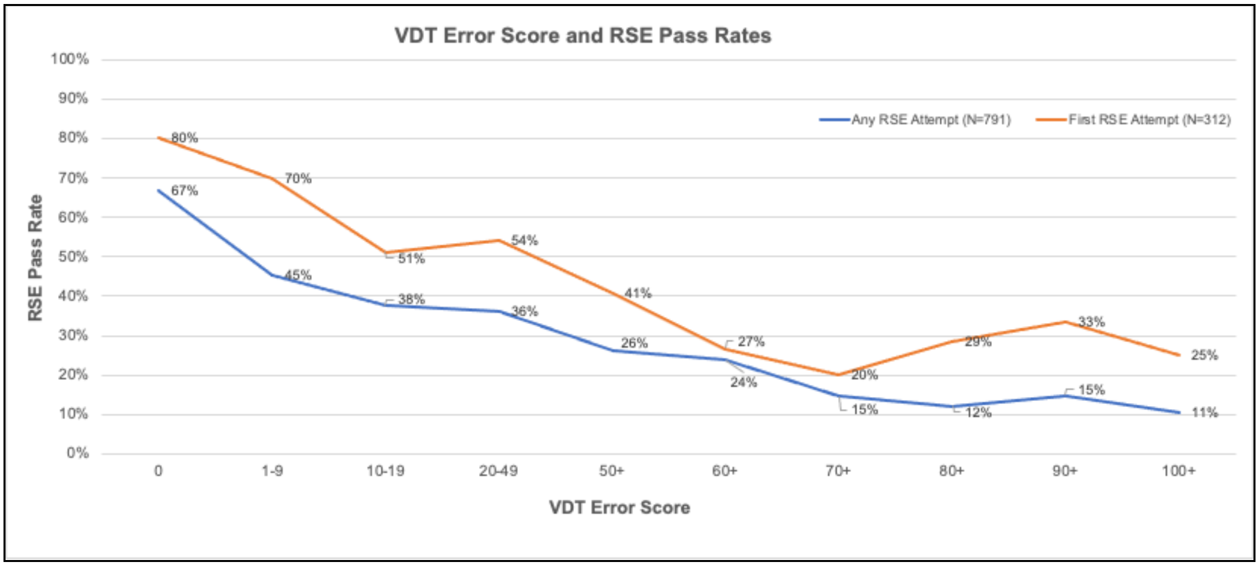

Figure 5: VDT Error Score and RSE Pass rates across all RSE attempts and the first RSE attempt only. Note: Larger VDT Error Scores indicate poorer driver performance.

<u>Significant errors and infractions made by customers during the VDT</u>:
Additionally, we examined the proportions of serious errors and infractions made by customers during the VDT. Table 9 below depicts these results among customers that failed the RSE and customers that passed the RSE.

Table 9: Significant driver errors made by customers during the VDT and the relative odds of failing the RSE. For example, a relative odds value of '1.85' implies an '85%' increased likelihood of failing the RSE if the VDT error was made vs. not made. Note: Green highlighted cells indicate statistically significant values and yellow highlighted cells indicate borderline statistically significant values.

	Any RSE Attempt (N=791)			**First RSE Attempt (N=312)**		
Significant driver error in VDT	% Who Made Error Within RSE=Fail	% Who Made Error Within RSE=Pass	Relative Odds of Failing RSE if Error was Made (95% CI)	% Who Made Error Within RSE=Fail	% Who Made Error Within RSE=Pass	Relative Odds of Failing RSE if Error was Made (95% CI)
Crashing into pedestrians	10%	8%	1.19 (0.71-1.97)	10%	7%	1.38 (0.62-3.09)
Crashing into vehicles	42%	29%	1.76 (1.29-2.38)	39%	25%	1.91 (1.18-3.11)
Crashing into objects	45%	41%	1.14 (0.86-1.53)	48%	42%	1.26 (0.80-1.97)
Stop sign	34%	28%	1.34 (0.98-1.84)	33%	25%	1.52 (0.92-2.49)
Red light	30%	22%	1.50 (1.08-2.10)	34%	23%	1.72 (1.04-2.84)
Navigation	22%	13%	1.90 (1.28-2.83)	20%	10%	2.28 (1.20-4.34)
Unsafe following distance	15%	16%	0.93 (0.63-1.37)	19%	15%	1.25 (0.69-2.26)
School zone violation	30%	27%	1.13 (0.82-1.55)	29%	28%	1.07 (0.65-1.76)
Pedestrian crosswalk	16%	15%	1.05 (0.70-1.56)	14%	16%	0.90 (0.48-1.69)
School bus violation	7%	4%	1.72 (0.90-3.31)	6%	2%	2.80 (0.83-9.5)

Emergency vehicle violation	64%	62%	1.05 (0.78-1.41)	65%	66%	0.98 (0.61-1.57)
Yielding to vehicles	43%	44%	1.00 (0.75-1.33)	43%	44%	0.95 (0.60-1.49)
Construction zone violation	30%	23%	1.48 (1.06-2.05)	26%	21%	1.33 (0.78-2.24)
Too slow for conditions	33%	30%	1.15 (0.84-1.57)	32%	31%	1.08 (0.66-1.75)
Too fast for conditions	31%	23%	1.51 (1.08-2.09)	27%	24%	1.17 (0.70-1.96)
Jerky braking	295	23%	1.34 (1.02-1.87)	31%	26%	1.25 (0.76-2.06)
Oversteering	51%	46%	1.19 (0.90-1.59)	47%	42%	1.22 (0.78-1.92)
Lane weaving	59%	52%	1.33 (1.01-1.77)	61%	52%	1.44 (0.91-2.27)
Driving off the road	53%	40%	1.70 (1.27-2.27)	47%	41%	1.31 (0.83-2.05)
Driving off center of lane	65%	55%	1.54 (1.15-2.06)	70%	53%	2.10 (1.31-3.37)
Not scanning for hidden hazards	67%	57%	1.51 (1.13-2.03)	65%	57%	1.38 (0.87-2.19)

Compared to customers who passed the RSE, customers who failed the RSE were more likely to have the following errors:

- Collide with vehicles
- Drive through a red light
- Drive through a stop sign
- Not follow the navigational directions
- Not yield to a stopped school bus
- Speed in a construction zone
- Aggressively speed along the driving route
- Exhibit jerky braking behaviors
- Weave in the roadway lane
- Oversteer
- Drive off the road
- Drive off the center of the roadway lane
- Not scan for hidden hazards

IMPLICATIONS OF RESULTS AND RECOMMENDATIONS FOR NEXT STEPS

The results from the pilot established foundational evidence across success criteria pertaining to feasibility of integration, low frequency of simulator-based motion sickness, and initial efficacy demonstrating that VDT performance was predictive of RSE results. These results set the stage for the VDT as a tool to improve the educational and testing ecosystem for all residents seeking a license, for driving schools preparing students for the RSE, and for MPI administrators and examiners.

<u>Implications</u>:

Consider that the VDT:

- **Works in limited resource settings.** The VDT is designed to operate in locations where space is at a premium, available staff resources are limited, and other important services need to be provided. This bodes well for any future use of the VDT by MPI – in its Centres or in other outposts that are easily accessible by customers.

- **Is a powerful and efficient way to deliver credible feedback to customers regarding their safe driving skills at a critical safety-intervention opportunity - the point of licensure.** The results indicated that customers found the VDT realistic and relevant, and customers were curious to know how they performed in the VDT. If feedback were to be enabled in a future deployment, MPI would be positioned to provide this feedback to all its RSE customers, along with detailed information about areas for improvement and a personalized intervention grounded in evidence and tailored to the specific improvement needs of each customer (Figure 6).

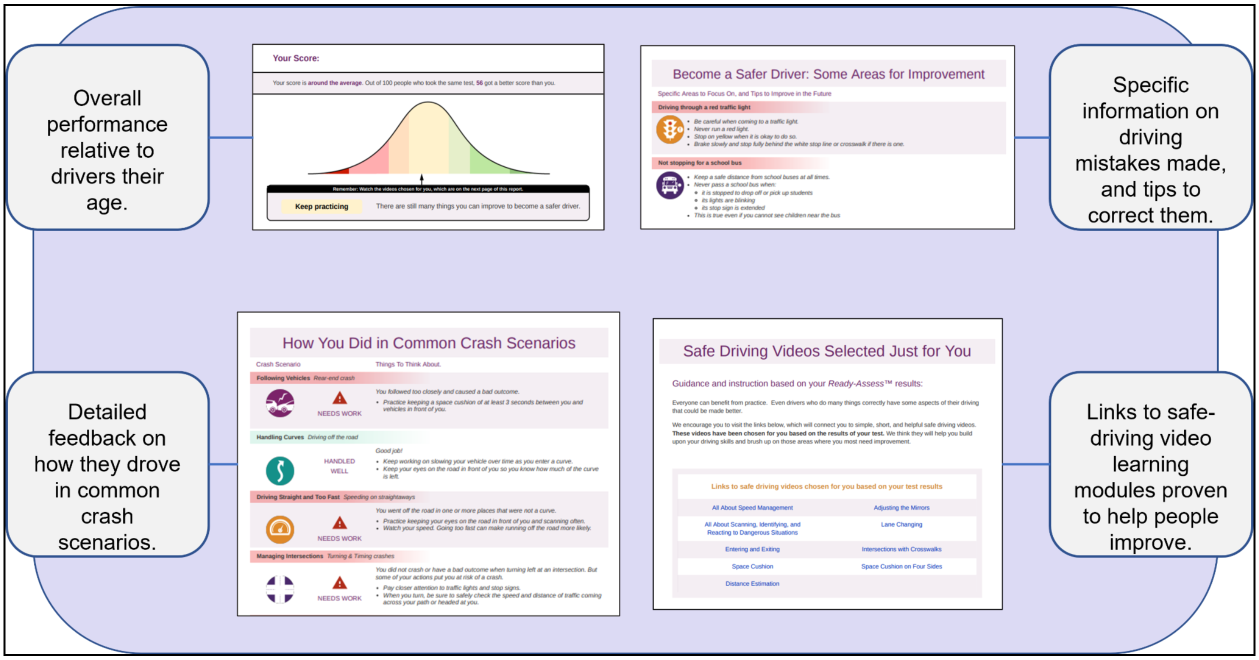

Figure 6: Overview of personalized feedback automatically delivered by the VDT. Full sample reports can be viewed here: https://tinyurl.com/samplefeedbackreports

- **Can serve as a lens through which MPI can better understand how Manitobans drive.** Each virtual driving test produces hundreds of driving-specific metrics, which in turn, can fuel a database of driving performance and common errors that cannot realistically be gathered in any other way. Such a robust and regularly updated database of driving behaviors would make for a powerful tool to aid in policy decisions, curriculum development, benchmarking, and oversight efforts.

- **Is predictive in a way that can deliver practical benefits to a variety of stakeholders.** The demonstrated strong relationship between VDT performance and RSE outcomes opens the door to integrating the VDT more formally into the current licensing workflow. Using the VDT to determine RSE readiness, or even to prioritize RSE scheduling:
  - Would directly assist MPI in its efforts to manage volume in its Centres and keep its examiners safe.
  - Should benefit driving schools, who would be well armed to provide students (and students' parents) with a data-driven assessment of whether they are ready to schedule an RSE.
  - Provide prepared MPI customers with feedback to be safer during the independent driving phase and improved RSE scheduling, and underprepared customers with the information they need to continue preparing themselves for the RSE and for independent driving.

- **Has the potential to serve the province in a *variety* of important ways.** The focus of the DTAP program was specifically on the VDT, the RSE workflow, and its function within Service Centres that administer the RSE. This is only one of many potentially important applications; however, the VDT can be applied towards several types of practical use cases including:
  - *The traffic court system*. Traffic citations present a powerful 'teachable moment' opportunity to drivers of all ages. Committing a traffic violation can be a sign that someone is underprepared to drive safely, and unsafe driving presents a serious public safety issue.

    Several judges, prosecutors, and other officials who encounter at-risk drivers are currently using the VDT – providing individuals who have committed non-criminal traffic violations with credible feedback, containing an evidence-based intervention.

  - *As part of driver's education and behind-the-wheel training*. The VDT is also used by driving schools. It is a complement to important on-road instruction, and it provides instructors (and the parents of learner-drivers) with information that they otherwise cannot get in any reliable and consistent way. Specifically, the VDT allows schools to safely expose their students to dangerous scenarios <u>and</u> provides them with a valuable glimpse into how a student will drive when no one else is in the vehicle with them.

  - *A tool to strengthen current employee driver training initiatives*. Organizations with drivers and vehicle fleets are using the VDT to provide employees with helpful feedback, to pinpoint common errors made by company drivers, and to continually update and improve internal safety training programs.

  - *A way to pre-screen driver employees*. Organizations deploy the VDT as part of the hiring process for individuals who will be required to drive corporate vehicles. In this way, organizations can reduce risk and increase overall safety by screening out applicants who are clearly not prepared to drive safely and avoid crashes.

Table 10 below depicts several commercial and non-commercial applications that the VDT is currently used in.

Table 10: Exemplary and current VDT applications and use cases across both commercial and research settings.

Setting	Context	VDT Application and Use Case	Key Partner(s)
Commercial	RSE workflow	Customer preparedness, safety screen for RSE	Ohio Bureau of Motor Vehicles, Manitoba

			Public Insurance Corporation, Delaware Division of Motor Vehicles
Commercial	Formal state-approved behind-the-wheel instruction	VDT integrated into state-approved BTW instruction as a point-assessment for novice drivers seeking licensure	Ohio Department of Public Safety
Commercial	Adolescent primary care	VDT as a point-of-care assessment+intervention during annual wellness visits	Children's Hospital of Philadelphia Primary Care Network
Commercial	Juvenile Traffic Court Systems	Dispositional and diversionary tool for non-criminal traffic offenses (fee-per-test model)	Various juvenile traffic court systems in Ohio and Nevada
Commercial	Adult and adolescent neuropsychological clinical practice	VDT as a point-of-care functional assessment when there is suspicion or formal diagnosis of neurocognitive impairment (fee-per-test model)	The Center for Neuropsychology and Counseling (Chalfont, PA)
Commercial	Consumer-based point-of-service assessment+feedback, particularly when formal behind-the-wheel training is not mandatory or not accessible	VDT installed at approved testing centers (e.g., local driving schools) to administer VDT to a private consumer base demanding this service (fee-per-test model)	Various DDI-approved testing centers in Pennsylvania and Florida
Commercial	Adult employee populations	Assessment+feedback for existing employee driver populations or as a pre-employment screen	Exelon Corporation (and various smaller employers)
Commercial	High school teen populations	Assessment+feedback for targeted outreach to high school teen populations	Center for Safe Alaskans (Anchorage, AK)
Research	Neurosurgery	VDT to collect functional performance data post-neurosurgery procedures among patients with traumatic brain injury	University of Pennsylvania (Philadelphia, PA)
Research	Adult HIV+ population	VDT to collect functional performance data for individuals with HIV-associated neurocognitive disorders (HAND)	Drexel University (Philadelphia, PA)
Research	Adolescents with developmental	VDT integrated into a magnetoencephalography (MEG)	Children's Hospital of Philadelphia

	disorders	protocol to task localized regions within the brain	(Philadelphia, PA)
Research	Adolescents	VDT to collect functional performance data among teens with executive function impairment	Children's Hospital of Philadelphia (Philadelphia, PA)

Recommendations:

The strong pilot results and their implications suggest that the VDT should be implemented as part of Manitoba's overall strategy to deliver high-quality driver's education and evaluate new drivers seeking independent driving privileges. Determining the most advantageous opportunities to do so will be key to long-term success (increased safety during the RSE, better prepared drivers, and long-term crash reduction). We recommend the following initial applications for a larger-scale deployment of the VDT:

1. **Allow all customers to self-assess and receive feedback using the VDT, without penalty.** Leverage the VDT to take advantage of the critical safety-intervention opportunity provided at the point of licensure. Customers who are unsure about their RSE readiness, can use the insights in their VDT results to voluntarily postpone their RSE. Additionally, offering the VDT as a self-assessment tool could be an effective way to address issues of equity and access that exist for people who live in more remote locations, or who otherwise lack the time or resources for formal behind-the-wheel instruction.

2. **Implement the VDT within driving schools, as a valuable new component of driver's education.** This would be a useful complement to both classroom and behind-the-wheel instruction as it would allow instructors to safely expose students to crash scenarios and better understand how students are likely to behave when driving on their own. It would also give these schools a powerful and standardized tool for assessing RSE readiness – a way to inform students (and their parents) about whether they are ready to attempt a RSE, and a scientifically valid way to advise their customers about the need for additional practice. Additionally, incorporating the VDT in the driving schools can be an effective strategy for MPI to build a large and robust database of driver readiness and common errors. As previously discussed, this would allow for enhanced program monitoring and updating as well as help to guide critical policy and curriculum decision making.

3. **Use the VDT as a safety-screen prior to customers attempting a RSE.** Leverage the objective and evidence-based results from the VDT to ensure that it is safe for a customer to attempt a RSE (or postpone to a future date). This would result in (1) a safer ecosystem for administering the RSE; (2) focusing limited examiner time on more prepared customers; and (3) helping to alleviate workflow burdens due to underprepared

customers. This application could be implemented in the Service Centre locations, within driving schools, or other public outposts that are accessible to customers.

4. **Use the VDT to assess medical fitness-to-drive.** Driving is cognitively demanding and an instrumental activity of daily living. New research demonstrates that the VDT can detect degraded driving performance among individuals with neurocognitive impairment. Additionally, the VDT provides objective feedback to both the driver, their families, and to official decision makers. In these cases, the VDT can be used in lieu of currently utilized in-office assessments and on-road cognitive evaluations.

LIMITATIONS

Our results are not without limitations. We could not leverage the full dataset (N=967) in the efficacy analysis as nearly 18% of the data either did not fully complete the VDT (to generate a VDT Error Score) or did not have a corresponding RSE result after completing the VDT. We did not document the reasons for why the 144 customers (15%) did not fully complete the VDT (this was due to practical and logistical reasons as customers attempted the VDT without any MPI staff supervision or oversight). However, the overall completion rate of 85% significantly improved in the later months of the pilot (>90%) which may have been an indication of improved staff workflow and training procedures. Also, not completing the VDT was statistically associated with failing the RSE (also supported by previous work conducted in other jurisdictions outside of Manitoba). Additionally, we suspected there were some outlier customers (e.g., those who performed poorly in the VDT and passed the RSE); however, these were a minority of cases and the final analysis used all the data available (including suspected outliers). DDI did not receive any numeric scores or qualitative examiner comments from the RSE (only a Pass/Fail indication). Finally, the VDT required a basic ‘high-speed’ Internet connection to automatically process and score VDT data (minimum download speed ≥ 25mbps; minimum upload speed ≥ 3mbps). This was not an issue in the five Service Centre locations where the VDT was piloted (no loss of data was observed); however, this could be an issue in more remote Service Centre locations where poorer connectivity is an existing problem (Council of Canadian Academies, 2021). In these situations, workaround procedures to successfully implement the VDT can be utilized (e.g., boosting wifi speeds, hotspotting, or implementing the VDT where data is stored, processed, and scored locally).

REFERENCES

Council of Canadian Academies. (2021). Waiting to Connect: The expert panel of high-throughput networks for rural and remote communities in Canada. https://cca-reports.ca/wp-content/uploads/2022/01/Waiting-to-Connect_FINAL-EN_digital.pdf.

Curry, AE, Metzger, KB, Williams, AF, & Tefft, BC. (2017). Comparison of older and younger novice driver crash rates: Informing the need for extended Graduated Driver Licensing restrictions. Accident Analysis and Prevention, 108, 66–73.

Curry, AE, Pfeiffer, MR, Durbin, DR, Elliott, MR. (2015). Young driver crash rates by licensing age, driving experience, and license phase. Accident Analysis and Prevention, 80, 243–250. https://doi.org/10.1016/j.aap.2015.04.019

Grethlein, D, Pirrone, V, Devlin, KN, Dampier, W, Szep, S, Winston, FK, Ontanon, S, Walshe, EA, Malone, K, Tillman, S, Ances, BM, Kandadai, V, Kolson, DL, Wigdahl, B. (2022). Examining virtual driving test performance and its relationship to individuals with HIV-associated neurocognitive disorders. Front. Neuroscience., 24 August 2022. https://doi.org/10.3389/fnins.2022.912766.

Grethlein, D, Winston, FK, Walshe, EA, Tanner, S, Kandadai, V, Ontanon, S. (2020). Simulator Pre-Screening of Underprepared Drivers Prior to Licensing On-Road Examination: Clustering of Virtual Driving Test Time Series Data. J Med Internet Res. 22(6):e13995. https://pubmed.ncbi.nlm.nih.gov/32554384/.

Lee, J. (2021). Children's Hospital of Philadelphia launches an innovative virtual teen driving assessment program with a $4 million grant from NJM insurance group. PRNewswire. Available online at: https://www.prnewswire.com/news-releases/childrens-hospital-of-philadelphia-launches-innovative-virtual-teen-driving-assessment-program-with-4-million-grant-from-njm-insurance-group-301308005.html (accessed March 15, 2022).

McDonald, CC, Curry, AE, Kandadai, V, Sommers, MS, Winston, FK. (2014) Comparison of teen and adult driver crash scenarios in a nationally representative sample of serious crashes. Accident Analysis and Prevention, 72, 302–308. https://doi.org/10.1016/j.aap.2014.07.016

Mirman, JH, Curry, AE, Winston, FK, Wang, W, Elliott, MR, Schulteis, MT, Fisher-Thiel, MC, Durbin DR. (2014). Effect of the teen driving plan on the driving performance of teenagers before licensure: a randomized clinical trial. JAMA Pediatr. 168(8):764-71. doi: 10.1001/jamapediatrics.2014.252.

Tefft, B. (2017). Rates of Motor Vehicle Crashes, Injuries, and Deaths in Relation to Driver Age, United States, 2014 – 2015. AAA Foundation for Traffic Safety. https://aaafoundation.org/rates-motor-vehicle-crashes-injuries-deaths-relation-driver-age-494 united-states-2014-2015/

Transport Canada. (2020). Canadian Motor Vehicle Traffic Collision Statistics: 2020. Available online at: https://tc.canada.ca/en/road-transportation/statistics-data/canadian-motor-vehicle-traffic-collision-statistics-2020 (accessed October 5, 2022).

Walshe EA, Elliott, MR, Cheng, S, Curry, AE, Seacrist, T, Oppenheimer, N, Wyner, AJ, Grethlein, D, Gonzalez, AK, Winston, FK. (2022). Novel use of a virtual driving assessment to classify driver skill at the time of licensure. Transportation Research Part F: Psychology and Behavior. https://doi.org/10.1016/j.trf.2022.04.009.

Walshe EA, Romer, D, Kandadai, V, Winston, FK. (2020). A Novel Health-Transportation Partnership Paves The Road For Young Driver Safety Through Virtual Assessment. Health Affairs. 39(10). https://www.healthaffairs.org/doi/10.1377/hlthaff.2020.00802.

Winston FK, Kandadai V, Hill SD. (2019). Ohio Portable Driver Simulation System Pilot: Implementation of a Virtual Driving Test Within Ohio's Driver Licensing Workflow (Research Brief). Philadelphia, PA. Children's Hospital of Philadelphia Research Institute.