Detecting driver fatigue using heart rate variability: A systematic review

Accident Analysis and Prevention 178 (2022) 106830

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Detecting driver fatigue using heart rate variability: A systematic review

Ke Lu [a,b,*], Anna Sjörs Dahlman [a,b,c], Johan Karlsson [b,d], Stefan Candefjord [a,b]

[a] Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden
[b] SAFER Vehicle and Traffic Safety Centre, Chalmers University of Technology, Gothenburg, Sweden
[c] Swedish National Road and Transport Research Institute (VTI), Linköping, Sweden
[d] Autoliv Research, Autoliv Development AB, Vårgårda, Sweden

ARTICLE INFO

Keywords:
Heart rate variability
Fatigue
Sleepiness
Driver monitor

ABSTRACT

Driver fatigue detection systems have potential to improve road safety by preventing crashes and saving lives. Conventional driver monitoring systems based on driving performance and facial features may be challenged by the application of automated driving systems. This limitation could potentially be overcome by monitoring systems based on physiological measurements. Heart rate variability (HRV) is a physiological marker of interest for detecting driver fatigue that can be measured during real life driving. This systematic review investigates the relationship between HRV measures and driver fatigue, as well as the performance of HRV based fatigue detection systems. With the applied eligibility criteria, 18 articles were identified in this review. Inconsistent results can be found within the studies that investigated differences of HRV measures between alert and fatigued drivers. For studies that developed HRV based fatigue detection systems, the detection performance showed a large variation, where the detection accuracy ranged from 44% to 100%. The inconsistency and variation of the results can be caused by differences in several key aspects in the study designs. Progress in this field is needed to determine the relationship between HRV and different fatigue causal factors and its connection to driver performance. To be deployed, HRV-based fatigue detection systems need to be thoroughly tested in real life conditions with good coverage of relevant driving scenarios and a sufficient number of participants.

1. Introduction

Driver fatigue is a major concern for road safety, and it accounts for 10–30 % of all fatal crashes (Hallvig et al., 2014; Philip and Åkerstedt, 2006; Zwahlen et al., 2016). Therefore, driver fatigue detection systems could potentially reduce fatigue related road fatalities and severe injuries. In the European Union, driver monitoring systems will become mandatory for new produced vehicles (European Parliament and Council, 2019) and it will become part of Euro NCAP safety assessment (Euro NCAP, 2017).

Fatigue is a complex phenomenon caused by multiple factors, and there is no consensus in the literature on the definition of fatigue and its relationship to sleepiness (Weinbeer et al., 2018). These terms have often been used synonymously in the literature. In this review, we will not develop the definition of fatigue or sleepiness and distinguish between them. Instead, we will break them down into common causation factors for fatigue in road driving (Fig. 1), and both terms will be used with the intention to follow the original literature cited in the review. It has been suggested that driver fatigue has both sleep related and task related causes (May and Baldwin, 2009). As shown in Fig. 1, sleep related fatigue is influenced by the circadian rhythm of sleepiness as well as the sleep homeostat, which depends on sleep duration and time awake since the last sleep episode. Task related fatigue is influenced by the driving itself and depends on time on task and the mental task load. Both underload and overload can contribute to fatigue, and the influence on driver performance and countermeasures may vary accordingly (Williamson et al., 2011). It is worth noting that studies focusing on driver sleepiness can include not only sleep related factors but also task related factors, and sleepiness measures such as subjective sleepiness rating could be influenced by task related factors as well (Åkerstedt et al., 2014).

Current fatigue detection systems are typically based on assessments of either driving performance such as speed and steering, facial features such as head pose, eye closure, and eye gaze, or physiological measurements such as electroencephalography (EEG), electrocardiography (ECG) and electromyography (EMG). Most of the current commercially available systems are based on driving performance and facial features detected by cameras (Chowdhury et al., 2018). However, those methods

* Corresponding author.
E-mail address: ke.lu@chalmers.se (K. Lu).

https://doi.org/10.1016/j.aap.2022.106830
Received 1 November 2021; Received in revised form 5 July 2022; Accepted 1 September 2022
Available online 22 September 2022

will be challenged by the increasing application of vehicle automation systems. SAE international defines 6 levels of driving automation from Level 0 (no automation) to Level 5 (full automation) (SAE, 2016). Many currently produced cars are equipped with Level 1 and 2 automation systems. In this case, speed and steering could be controlled by the vehicle through lane keeping and adaptive cruise control functions, and that information can then no longer be used as a measure of the driver's performance (Gonçalves and Bengler, 2015). Prototypes of Level 3 and 4 vehicles are on trial in demonstration sites, in which case the driver will no longer be responsible for monitoring the environment when automation is active. When reaching Level 3 and above in the future, facial features including eye gaze, eyelid closure and head positioning may not be available as indicators of driver fatigue (Wörle et al., 2019). At the same time, fatigue could become more frequent under automated driving if the driver does not have active task engagement (Ahlström et al., 2021; Körber et al., 2015; Schömig et al., 2015; Vogelpohl et al., 2019). Physiological measurements of fatigue could become a potential solution to this challenge.

A recent review investigated the performance of driver sleepiness detection methods using physiological signals (Watling et al., 2021). They concluded that progress is needed to reach sufficient specificity and sensitivity and that using multiple physiological signals resulted in improved assessment. However, many physiological sensors are not favorable for daily usage due to the obtrusive measurement setups that require attachment of gel electrodes and wiring (Lohani et al., 2019). Heart rate variability (HRV) has drawn particular interest due to its relationship with fatigue and ease of measure in real life (Lohani et al., 2019). HRV is the fluctuation of time between adjacent heart beats. The variation of heart rate (HR) is generated by heart-brain interaction through the sympathetic and parasympathetic branches of the autonomic nervous system. HRV reflects the response of cardiac autonomic nerves to inputs from baro-, chemo-, nasopharyngeal and other receptors, as well as central autonomic commands that are associated with stress, physical activity, arousal, sleep, etc. (Silvani et al., 2016). Several sleep laboratory studies show that HRV can be a good indicator for vigilance state measured by reaction speed to visual stimulus under total sleep deprivation (Chua et al., 2012) and partial sleep deprivation (Henelius et al., 2014; Kaida et al., 2007). HRV has also been shown to be associated with cognitive task demand and time on task effects (Hidalgo-Muñoz et al., 2018; Luque-Casado et al., 2016). With the development of unobtrusive sensing techniques, HR and HRV could be measured through wearable sensors (Sikander and Anwar, 2019; Zheng et al., 2014) or vehicle integrated sensors (Leonhardt et al., 2018; Pinto et al., 2017) in daily driving scenarios. Several studies have approached the relationship between driver fatigue and HRV parameters by building fatigue classifiers based on HRV features (Abtahi et al., 2018; Buendia et al., 2019; Fujiwara et al., 2019; Kundinger et al., 2020a; Lenis et al., 2016; Li and Chung, 2013; Mahachandra et al., 2012; Patel et al., 2011; Persson et al., 2021; Zeng et al., 2020).

Although many studies have reported HRV as a driver fatigue indicator, there is not yet a consensus on how HRV changes during the development of driver fatigue. Several reviews on driver monitoring systems have included solutions based on HR (Sahayadhas et al., 2012; Sikander and Anwar, 2019; Watling et al., 2021), but the relation between HRV and fatigue has not been summarized. This review aims to summarize and analyze the literature on 1) how HRV features change under fatigue, 2) Performance of HRV based fatigue detection systems, and 3) the potential for HRV to be used as an indicator of driver fatigue in real life settings. We conducted a systematic review of studies that have explored the relationships between HRV and driver fatigue and that have developed HRV based driver fatigue detection systems.

2. Methods

2.1. Search methods

Three databases deemed most relevant for the research topic were searched in this systematic review, i.e., PubMed, Scopus, and Web of Science (Web of Science Core Collection). The search was conducted in July 2021 and there was no limit to the starting date. The terms used in the search were *'(heart rate OR hr OR hrv) AND (sleep* OR drows* OR fatigue) AND driver'*. The terms were searched for in the fields of title, abstract and keywords. Metadata (title, author list, journal, volume, etc.) of the articles from the search results together with the abstracts were downloaded and imported to Rayyan (Ouzzani et al., 2016) for screening and selection.

2.2. Eligibility criteria

We included only original research journal articles written in English. As we were aiming to investigate solely the relationship between HRV and driver fatigue, the included studies should report the relation between driver fatigue and HRV explicitly. Studies that mix HR or HRV together with other measurements were excluded from the review. Studies that were not conducted with car driving task, e.g., airplane, ship, train driving, as well as race car driving were also excluded. Since the focus was on mental fatigue, studies that targeted physical fatigue were also excluded. The selection process is shown in Fig. 2. In total, 977 records were found in the three databases and 633 records remained after duplication removal, in which 384 journal articles in English were kept for screening. After reading through the title and abstract, 348 articles were excluded, and 36 articles were left for full text assessment.

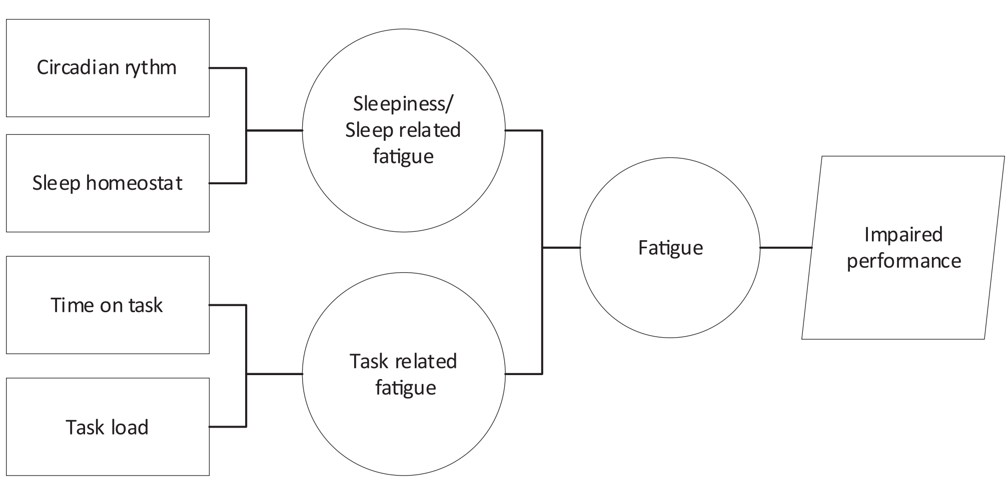

Fig. 1. Common factors that lead to drivers' mental fatigue.

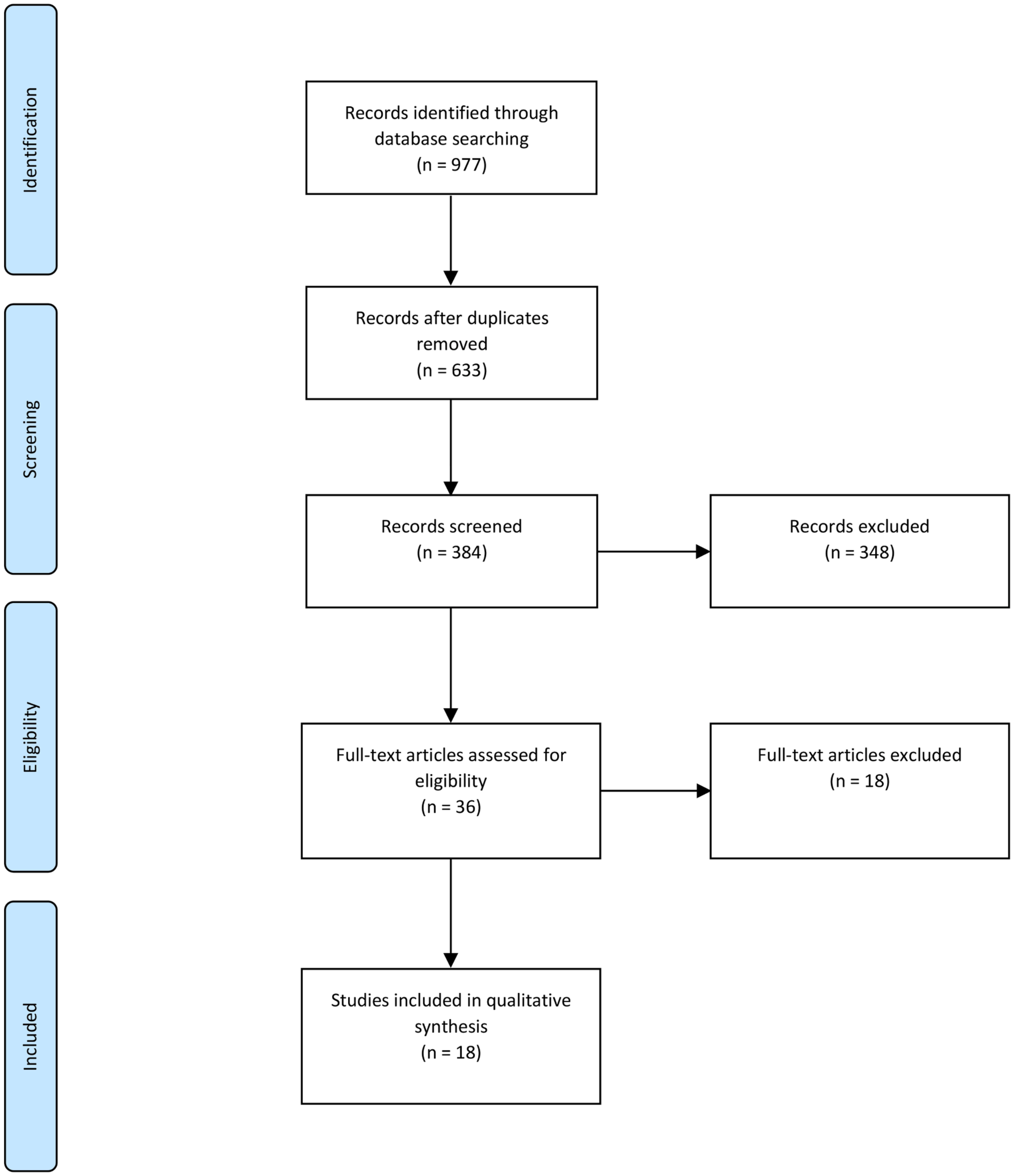

Fig. 2. Flow chart of article identification, screening, and selection.

Following full text investigation, 18 articles were removed, leaving 18 articles for review. Among the 18 removed articles, six articles were removed due to mixing HRV with other measurements, four studies were not performed with car driving tasks, two did not include HR or HRV, two did not have a reference fatigue measure, two used HR or HRV directly as a valid fatigue indicator, one demonstrated the development of a HR based fatigue detection system without evaluation, and one studied how a HR based fatigue detection system correlated to driver behavior. The eligibility criteria were decided by all authors together. Search, screening, and final article selection were performed by the first author (K.L).

2.3. Data extraction

The information extracted from the selected publications included demographics, driving tasks, measurement methods, classification approach, and detection performance and results. For demographics information, number of participants and the age of the participants were extracted. For the driving task, we extracted the type of driving task (simulator or on-road), the duration of each driving session, how many driving sessions each participant performed, and any manipulation method to introduce fatigue. Regarding measurements, we extracted the method for the HR measurement and how the reference level of driver fatigue was measured. For studies that aimed to build classifiers the validation methods and detection performance was extracted. For studies that reported HRV under fatigued conditions compared to alert conditions, we extracted the HRV features that differed between conditions and the direction of the changes. We focused on standard HRV features included in the Task Force guidelines for HRV measurements (Malik et al., 1996), which have been widely used in this field and in the included studies. Included HRV features and their short descriptions can be found in Table 1, and more detailed definition information can be found in (Malik et al., 1996; Shaffer and Ginsberg, 2017).

3. Results and discussion

Details of the reviewed studies are listed in Table 2. A substantial variation in the study implementation and results can be found across all studies. In total, 11 of the reviewed studies demonstrated differences in HRV features between fatigued states and alert states (Table 2). In studies that examined the HR level (or mean NN interval), reduced HR was found when drivers were fatigued, with only one exception where no significant change was found (Egelund, 1982). However, when it comes to the other time and frequency domain HRV features, the changes are not consistent across all studies. There were 11 reviewed studies that had developed HRV based fatigue detection systems and the reported detection performance ranged from 44 % to 100 % percent in accuracy. The difference in outcomes could be the result of differences in several aspects of the study designs including experiment setups, fatigue definition, and validation methods. The differences in study design also makes it difficult to compare the results quantitatively across all studies. In the following discussion we will highlight several key elements in the study design and their potential influences on the outcomes.

3.1. Study design

3.1.1. Study population

The sample sizes of reviewed studies ranged from 2 to 86. Half of the studies had relatively small study samples with less than 10 participants. There were six studies that included >30 participants in their experiments (Buendia et al., 2019; Fujiwara et al., 2019; Kundinger et al., 2020a, 2020b; Persson et al., 2021; Vicente et al., 2016).

Age has been known as an influencing factor for the majority of linear and nonlinear HRV indices for both long term (Voss et al., 2009) and short term (Voss et al., 2012) measurements. Among the reviewed articles, five studies included participants with a wide age range (Buendia et al., 2019; Kundinger et al., 2020a, 2020b; Patel et al., 2011; Persson et al., 2021). In addition, one study approached the influence of age by separating participants into two age groups (Kundinger et al., 2020b), the study suggests that a model developed with a specific age group is not well suited for another age group.

Sex is another demographic factor potentially associated with HRV. Sex differences in HRV measures have been reported by many studies (Koenig and Thayer, 2016). Among the reviewed articles, (Zeng et al., 2020) investigated sex differences for 13 HRV measures in both alert and fatigued states. In that study, more measures in the fatigued state showed significant differences between sexes than in the alert state, and male drivers had more measures with significant differences between fatigued and alert states than female drivers.

Table 1
Summary of HRV features that are included in the review.

Domain	Label	Feature
Time	HR	Mean heart rate value
	SDNN	Standard deviation of NN intervals
	RMSSD	Root mean square of successive differences
	pNN50	Percentage of intervals that differ > 50 ms from previous interval
Frequency	VLF	Very low frequency spectral power (0.0033–0.04 Hz)
	LF	Low frequency spectral power (0.04–0.15 Hz)
	LFnu	Normalized power in LF band
	HF	High frequency spectral power (0.15–0.4 Hz)
	HFnu	Normalized power in HF band
	LF/HF	Ratio between LF and HF

3.1.2. Driving task

About two thirds of the reviewed studies were performed in a simulator, and six studies used a real road driving task, (Buendia et al., 2019; Egelund, 1982; Jung et al., 2014; Persson et al., 2021; Salvati et al., 2021; Wang et al., 2019). The study by (Vicente et al., 2016) used data from both simulator and on-road driving. Two reviewed studies had an automated driving task with simulated SAE Level 2 driving (Kundinger et al., 2020a, 2020b), whereas remaining articles were performed with manual driving.

3.1.2.1. Fatigue manipulation. The reviewed studies have taken different approaches to introduce fatigue to the subjects. Circadian rhythms and sleep homeostasis are two main contributors to sleep related fatigue (Franken and Dijk, 2009). Six studies manipulated fatigue by letting the participants perform driving tasks at different times of the day (Buendia et al., 2019; Kundinger et al., 2020a, 2020b; Lee et al., 2019; Murugan et al., 2020; Persson et al., 2021). (Kundinger et al., 2020a; Vicente et al., 2016) introduced partial or complete sleep deprivation before the driving session. When it comes to task related fatigue, under-stimulated and prolonged driving is known to introduce higher risk of driver fatigue (Williamson et al., 2011). Several studies opted to use monotonous driving tasks to speed up the development of fatigue (Fujiwara et al., 2019; Kundinger et al., 2020a, 2020b; Lenis et al., 2016; Murugan et al., 2020). The duration of the driving task varied from 10 min up to several hours. For studies with continuous and prolonged driving tasks, the time-on-task also became an important factor for fatigue development.

3.1.3. Measurements

3.1.3.1. HR measurement method. Several types of HR measurement devices have been used in the studies included in the review. Conventional ECG with gel electrodes was used by the majority (Buendia et al., 2019; Egelund, 1982; Fujiwara et al., 2019; Lenis et al., 2016; Murugan et al., 2020; Patel et al., 2011; Persson et al., 2021; Vicente et al., 2016; Wang et al., 2019; Zeng et al., 2020). Wearable ECG-based HR chest straps is another solution that brings better usability than the ECG with gel electrodes, which was used by (Khamis et al., 2016) and (Lee et al., 2019). (Jung et al., 2014) used integrated ECG electrodes on the steering wheel. Photoplethysmography (PPG) based solutions can be even easier to use since they can be integrated in wrist bands (Kundinger et al., 2020a, 2020b; Lee et al., 2015) or the steering wheel (Rahim et al., 2015). This advantage may enable pervasive usage in daily driving scenarios for HRV based monitoring. However, PPG based solutions are more sensitive to motion artifacts compared to ECG based devices. Two studies (Kundinger et al., 2020a; Lee et al., 2015) compared the wrist band PPG to ECG for fatigue detection and show that PPG-based HR can be used for this application but with reduced detection performance compared to ECG. Salvati et al. (2021) used a microphone sensor integrated in the seat cover for HR detection.

3.1.3.2. Reference fatigue measure. Different approaches were taken to measure the fatigue level as the ground truth. Some studies provided insufficient information about the definition of fatigued state used. This makes it difficult to interpret results and compare results across different studies. Observer rating was the most used method for reference fatigue (Kundinger et al., 2020a, 2020b; Lee et al., 2019; Lenis et al., 2016; Murugan et al., 2020; Rahim et al., 2015; Vicente et al., 2016; Zeng et al., 2020). However, low inter-rater agreement has been found for observer ratings of fatigue (Ahlstrom et al., 2015). Subjective ratings were also used for many studies. The Karolinska Sleepiness Scale (KSS) is a well validated scale for subject rating (Kaida et al., 2006), it was used by several studies (Buendia et al., 2019; Khamis et al., 2016; Kundinger et al., 2020b; Persson et al., 2021; Salvati et al., 2021). Some studies used self-defined scales, e.g., (Wang et al., 2019) used a 4-level scale that

Table 2
Details of reviewed studies.

Study	Driving setting	Participants	Measurements	HRV response	Fatigue Detection System
(Salvati et al., 2021)	- Real road - Motor way 46.2 km, twice per drive in afternoon, 14 sessions for all	- 3, age > 30	- HR: capacitive microphone sensor, seat cover integrated - Fatigue: PERCLOS and Subjective rating (KSS), every 15 min		- Window: 5 min - Type: Rule based classification - Performance: Acc: 63 %
(Persson et al., 2021)	- Real road - Manipulation: daytime ● Set 1: motor way, 90 min*2 (day, night) ● Set 2: motor way, 135 min*2 (day, night) ● Set 3: rural road, 90 min*3 (day, evening, night)	● Set 1: 18 (M10, F8), mean age 41 ● Set 2: 24 (M12, F12), mean age 35 ● Set 2: 44 (M23; F21), mean age 44	- HR: ECG - Fatigue: Subjective rating (KSS), every 5 min	HR – SDNN + RMSSD + pNN50 + VLF n.s. LF + LFnu + HF n.s. HFnu – LF/HF +	- Window: 5 min - Type: Three class / binary classification Personal baseline: segments with KSS less than 5 - Validation: 10-fold, LOSO - Performance: Acc: 44 %, Sen: 33 %, Spe: 66 % (three-class, LOSO)
(Murugan et al., 2020)	- Simulator - Manipulation: daytime, secondary task - Monotonous speedway, 120 min	- 10 (M9, F1), age: 19–35	- HR: ECG - Fatigue: Video based labeling (Criteria not specified), multiple class labeling by normal, drowsy, fatigue, visual inattention, cognitive inattention		- Window: Not specified - Type: Five class / binary classification - Validation: Hold out 75 % training, 25 % testing - Performance: Acc: 58 % (five-class), Acc: 100 % (binary, normal-drowsy), Acc: 96 % (binary, normal-fatigue)
(Kundinger et al., 2020a)	- Simulator - Partial automated driving (ACC, acceleration and break, no take over) - Manipulation: daytime - Monotonous closed-loop track, highway with little traffic, 45 min*1 (morning/afternoon/evening)	- 15 young (M9, F6), mean age: 22.87 ± 1.81 - 15 older (M7, F8), mean age 67.60 ± 1.88	- HR: PPG wristband, ECG - Fatigue: Observer rating on video and micro-sleep by eye closure		- Window: 5 min sliding with 2 s increment - Type: Binary classification - Validation: 10-fold, LOSO - Performance: Acc: 73 %, (LOSO, PPG) Acc: 79 %, (LOSO, ECG)
(Kundinger et al., 2020b)	- Simulator ● Set 1: Manipulation: sleep deprivation - Monotonous closed-loop track, 60 min (sleep deprived and 30 min (normal sleep) ● Set 2: Manipulation: daytime - Monotonous closed-loop track, highway with little traffic, 45 min*1 (morning/afternoon/evening)	● Set 1: 10 (M9, F1), mean age: 24 ± 2.05 ● Set 2: 15 young (M9, F6), mean age: 22.87 ± 1.81 -15 older (M7, F8), mean age 67.60 ± 1.88	● Set 1: - HR: PPG, wristband - Fatigue: Observer rating on video and subjective rating (KSS every 10 min) ● Set 2: - HR: PPG, wristband - Fatigue: Subjective rating (KSS), every 5 min		- Window: 2 min sliding with 2 s increment - Type: Binary classification - Validation: 10-fold - Performance: Acc: 99 %
(Zeng et al., 2020)	- Simulator - Highway, 60 min, between 9:30 to 17.30	- 20 (11 M, F9), mean age 25.95 ± 2.67	- HR: ECG - Fatigue: Video based facial expressions, three levels	HR – SDNN+ RMSSD+ pNN50 + VLF + LF + LFnu n.s. HF + HFnu n.s.	
(Fujiwara et al., 2019)	- Simulator - Monotonous highway loop, 90 min*2 (11:00, resting and lunch in between)	- 34 (M25, F9) mean age 22.7 (18–36)	- HR: ECG - Fatigue: N1 onset, EEG-based sleep scoring		- Window: 3 min - Type: Anormal detection - Personal baseline: 0.12–1.5 h awake episodes Performance: 12/13 pre-n1 episodes were detected prior to sleep onsets, 1.7–times false positive per hour
(Buendia et al., 2019)	- Real road - Manipulation: daytime ● Set 1: rural road, 90 min*3 (day, evening, night) ● Set 2: rural road 90 min*3 (day, evening, night) ● Set 3: motor way, 135 min*2 (day, night)	● Set 1: 21 (M11, F10), mean age 44.8 ± 7.8 ● Set 2: 22 (M22, F11), mean age 45 ± 8.2 ● Set 2: 38 (M19; F19), mean age 35 ± 9.6	- HR: ECG - Fatigue: Subjective rating (KSS), every 5 min	HR – SDNN + RMSSD + LF + LFnu + HF + HFnu n.s. LF/HF +	
(Lee et al., 2019)	- Simulator - Manipulation: daytime - Euro Truck Simulator 2	- 6, age: 25–35	- HR: ECG chest strap and PPG wristband - Fatigue: Video of face and behavior, criteria not specified		- Window: 2 min - Type: Binary classification - Drowsy samples in the morning

(continued on next page)

Table 2 (*continued*)

Study	Driving setting	Participants	Measurements	HRV response	Fatigue Detection System
	- In total, 6 trails in morning and 16 after dinner or lunch				and awake samples after lunch or dinner are dropped - Validation: 10-fold - Performance: Acc: 70 % for ECG, Acc: 64 % for PPG
(Wang et al., 2019)	- Real road - Freeway, 345 km, (between 8:00–18:00)	- 10, mean age 32.6 (24–45)	- HR: ECG - Fatigue: Subjective rating (alert/slight fatigue/serious fatigue/drowsiness), every 15 min		- Window: sample by sample - Type: Anormal detection - Personal baseline: Entire driving session - Performance: All cases recognized, with less than 5 min delay
(Vicente et al., 2016)	● Set1: Simulator - Highway, 120 min - Manipulation: sleep deprivation ● Set2: Simulator - Highway, 100 min - Manipulation: sleep deprivation ● Set3: Real Road - Highway or road during 8 h working day	● Set1: 9 (M4, F5)) ● Set2: 11 (M5, F6) ● Set3:10 (M8, F2)	- HR: ECG - Fatigue: External observer (awake/ fatigued/ drowsy) for set 1 and 3, Sleep condition (not/partial/- sleep deprived) for set 1 and 2	LF/HF –	● Drowsiness episodes (set 1 & 3): - Window: 1 min - Type: Binary classification - Personal baseline: segments in first 3 min - Validation: LOSO - Performance: Sen: 59 %, Spe: 98 % ● Sleep-deprivation (Set 1, 2, & 3): - Window: 3 min - Type: Binary classification - Performance: Sen: 62 %, Spe: 88 %
(Lenis et al., 2016)	- Simulator - Two-lane street, monotonous, 40 min*7, between 1 am to 8 am	- 14 young (age and gender are not specified)	- HR: ECG - Fatigue: video observation for microsleep events	HR – pNN50 + RMSSD + SDNN +	
(Khamis et al., 2016)	- Simulator - Highway, 60 min *2 (with/ without vibration), during 9–10 am	- 3, age: 27–40	HR: ECG chest strap Fatigue: Subjective rating (KSS), before, during (after 30 min) and after driving, Driving performance (variation of lane deviation)	HR –	
(Rahim et al., 2015)	- Simulator - 2 h	- 2 (M1, F1)	- HR: PPG on steering wheel - Fatigue: Video of driver's behavior, criteria not specified	LF/HF –	
(Jung et al., 2014)	- Real road - 2 h	- 2 (M2, F0), age 27–31	- HR: ECG steering wheel integrated - Fatigue: driver's behavior (normal, drowsy, fatigued)	HR – SDNN – RMSSD – pNN50 – LF – HF – LF/HF +	
(Li and Chung, 2013)	- Simulator, - 10 min*2(alert/drowsy)	- 4 (M3, F1)	- HR: PPG - Fatigue: PERCLOS (alert (0 %-30 %)/ drowsy (30 %-40 %))	LF/HF +	- Window: 1 min - Type: Binary classification - Validation: LOO (sample) - Performance: Acc: 95 %, Spe: 95 %, Sen: 95 %
(Patel et al., 2011)	- Simulator - Manipulation: 2 h less sleep night before	- 12, mean age 47 ± 11	- HR: ECG - Fatigue: not specified	LF/HF –	- Type: Binary classification - Performance: Acc: 90 %
(Egelund, 1982)	- Real road - Highway, 340 km, between 1.30 pm and 5.30 pm	- 9 (M5, F3), age 19–22	- HR: ECG - Fatigue: driving distance	HR n.s. SDNN n.s. LF + (0.05–0.15 Hz)	

- Fields in HRV response and fatigue detection system left empty when it was not investigated by the study.
- For descriptions of driving scenarios, original phrasing from the reviewed article is used.
- For HRV response, '+' and '-' stands for higher and lower value under fatigue state comparing to alert state, respectively, 'n.s.' for no significant change.
- M = male, F = female, HR = Heart rate, ECG = electrocardiogram. PPG = photoplethysmography, EEG = electroencephalography, PERCLOS = percentage of eyelid closure, Acc = accuracy, Sen = sensitivity, Spe = specificity, LOSO = leave one subject out, LOO = leave one out.

has not been evaluated and the cut off level for classification was not reported. Another approach was to define fatigue based on percentage of eyelid closure over the pupil over time (PERCLOS) (Li and Chung, 2013). PERCLOS has been used extensively as a measure of fatigue but the relationship with subjective sleepiness is not straightforward (Sommer and Golz, 2010). (Fujiwara et al., 2019) used EEG signals to find the N1 sleep stage onset defined by alpha wave attenuation. One study did not use a reference measure for fatigue but used the driving distance or time-on-task as the reference (Egelund, 1982).

3.1.4. *Fatigue detection*

3.1.4.1. Time scale. Different temporal scales were used for analysis and for developing fatigue detection systems. Most studies targeted continuous analysis and detection, where data were segmented over the entire drive using moving or sliding window. The window length varied between 1 min and 5 min. Few studies used only one sample (a short driving session or a short window picked from the driving session) for each participant under each condition (fatigued/alert).

3.1.4.2. Learning and validation. Most of the reviewed detection systems were built with supervised learning methods, where each sample (usually containing data measured within a certain time window) used for the model training was labeled with the fatigue condition. (Fujiwara et al., 2019) and (Wang et al., 2019) applied semi-supervised and unsupervised learning with anomaly detection approaches instead. In these cases, the models were first built with the data under alert conditions or the entire dataset with the majority being alert conditions, and then the models were used to detect anomalies in data, which were identified as the fatigued conditions.

Validation methods can have a significant impact on the performance measure (Saeb et al., 2017). It should be taken into consideration that HRV measures have inter-individual differences in both resting level and response to different stimulations (Nunan et al., 2010; Ohyama et al., 2007). Data samples from the same driving session and participants are likely to be highly correlated. To evaluate real life performance for new users, the data from the same driving session and same participant needs to be separated from the training and test set. Some reviewed studies have applied leave one subject out (LOSO) cross validation where such separation was achieved (Kundinger et al., 2020a; Persson et al., 2021; Vicente et al., 2016). Most of the remaining studies used 10-fold cross validation or hold out validation without arrangement for participant separation in training and testing, which may bias the results and exaggerate detection performance in relation to future use in real life driving scenarios.

To deal with the inter-individual differences, some reviewed studies have applied personalization methods. (Vicente et al., 2016) and (Persson et al., 2021) used personalized baselines to create a personalized feature set that accounts for the basal level of personal HRV measures. For models developed using the anomaly detection approach (Fujiwara et al., 2019; Wang et al., 2019), each person has their own model based on data from themselves.

3.2. HRV response for driver fatigue

Not all reviewed studies reported how HRV variables were related to fatigue, i.e., the direction of change when going from alert to fatigued. In total, 11 studies reported the difference in HR or HRV between fatigued and alert states. Among all measures, the LF/HF and HR level (or mean NN interval) were most investigated (Table 2). Decreased HR (increased mean NN interval) in the fatigued state was reported by five out of six reviewed studies that investigated HR change, while the remaining study did not find a significant change. For other time and frequency domain HRV parameters, contradictory results can be found where both increased and decreased values have been reported (Table 2). Several reviewed studies (Li and Chung, 2013; Patel et al., 2011; Rahim et al., 2015; Vicente et al., 2016) and other studies (Awais et al., 2014; Byeon et al., 2006) have considered LF/HF to be an important indicator of fatigue, as a reflection of the balance between parasympathetic and sympathetic nerve activity. However, the changes of LF/HF were not consistent across all reviewed studies. The inconsistency can be caused by different experiment setups, including the driving task, cause of fatigue and level of fatigue. Small study samples could also limit the generalizability of some studies.

It has been hypothesized that fatigue activates the parasympathetic nervous system, which leads to higher levels of HF, whereas when the sleep demand is counteracted by subjects fighting to stay awake this will lead to sympathetic activation that increases LF (Vicente et al., 2016). Therefore, for real road driving, drivers might have higher intention to fight against sleepiness than in the simulator studies, which leads to higher sympathetic activation. However, the physiological base of such an assumption is questionable. The HRV LF power is reflecting a mixture of sympathetic and parasympathetic activities together with other factors and the LF power is thus not directly correlated to sympathetic nerve activity (Moak et al., 2007; Piccirillo et al., 2009). The physiological base for LF/HF is indistinct and to interpret LF/HF as the balance between the parasympathetic and sympathetic nerve activity has been challenged (Billman, 2013).

Differences in the cause of fatigue can also be the reason for different HRV responses in relation to fatigue. In sleep research, several studies have reported increased HRV measured as SDNN (Kaida et al., 2007) in sleepy subjects. Increased VLF and LF power (Henelius et al., 2014), (Chua et al., 2012) have been associated to decreased vigilance caused by total or partial sleep deprivaiton. In those studies the effect of sleep homeostats and circardian effects are involved. While falling asleep, reduced HRV has been observed (Shinar et al., 2006). For task related factors, studies have shown that HRV changes are reflecting the cognitive task demand (Luque-Casado et al., 2016) or the time on task effect. For the time on task effect, both increased (Matuz et al., 2021) and decreased (Luque-Casado et al., 2016; Melo et al., 2017) HRV has been reported. The difference could be caused by different task demands and engagement (Pendleton et al., 2016).

Even when studying the same type of fatigue, the level of fatigue is another factor that needs to be considered. In study by (Henelius et al., 2014), a strong correlation between HRV spectral power and the psychomotor vigilance performance was only found for high levels of sleepiness under sleep deprivation, but not slight vigilance decrement in an ordinary day. Hence, studies that target low levels of fatigue may have different results than those targeting high levels of fatigue.

3.3. Performance of fatigue detection

The performance measure of HRV based driver fatigue detection systems varied from poor to perfect across the reviewed studies. Very high accuracy (i.e., both high sensitivity and specificity) was found in some simulator studies without subject-wise separation in learning and testing. Due to the differences in study design, those performance measures have different meanings in practice and cannot be compared directly by the numbers. Several design aspects can have a significant impact on the study outcome. The factors that were discussed in the previous section, including how fatigue was introduced and measured, can also affect the performance of a fatigue detection system. Another key factor is whether the data from the same participant were separated from the training and testing set. The difference in detection performance measures between LOSO cross validation and k-fold cross validation was shown by (Kundinger et al., 2020a) and (Persson et al., 2021). Having real road scenarios rather than simulator, having a broader population coverage (e.g., sex, age, fitness, etc.), and using less accurate HR sensors can also lead to lower performance measures (Persson et al., 2021).

3.4. Use of HRV based detection in real life

A majority of the reviewed studies were conducted in a simulator environment, and few were conducted with controlled real road scenarios. Findings from those studies may face many challenges in real life scenarios due to a variety of contextual factors that can alter HRV. HR and HRV have been shown to reflect cognitive workload (Luque-Casado et al., 2016). In real life scenarios, varying complexity of the driving context and involvement of secondary tasks will influence HR and HRV as well as fatigue development. Changes in emotional states and stress, food intake, and change in environmental factors (e.g., altitude, temperature) can also introduce variation in HR and HRV (Appelhans and Luecken, 2006; Boos et al., 2017; Castaldo et al., 2015; Sollers et al., 2002). Whether those stimulations can overshadow the HRV changes due to fatigue needs to be investigated in real life driving scenarios. HRV based monitoring can also be challenging in people with certain medical conditions and medications whose HRV regulation is affected.

Having personalized algorithms could be a key element for an accurate HRV based fatigue detection system. Current studies involve measurements for each participant during one or two days only. However, the personal basal HRV level can change overtime and good strategies for personalization need to be investigated. At the same time, those personalization methods should also consider local regulations regarding personal data usage in practice.

3.5. Future perspectives

The development of driver fatigue is a complex process influenced by multiple factors, and so is the physiological regulation of HRV. With this review we are not able to conclude that there are solid relationships between HRV and driver fatigue. How HRV is related to different causes of fatigue and its relation to driving performance under different types of fatigue needs to be further investigated. It will be helpful for future studies to have a transparent reporting on factors that influence fatigue and on reference measures for fatigue.

There is a need to find a good reference measure for fatigue that can reflect deterioration of driving performance and safety outcomes. Some reviewed studies have used fatigue measurements that have not been validated. Finding a valid and reliable ground truth measure of fatigue is challenging. There are drawbacks with both the subjective and objective physiological measures of fatigue and sleepiness. The drivers might not be fully aware of or might not acknowledge their signs of fatigue. They can also experience difficulties in reporting the correct level on subjective rating scales. Objective physiological measurements often do not correspond fully to the level of fatigue or sleepiness experienced by the subject. At the same time, test procedures in newly developed regulations such as European Union General Safety Regulation could be used as a base for future study design (European Parliament and Council, 2019).

A recent on-road experiment with a relatively large population did not achieve a satisfactory result for fatigue detection with direct usage of HRV (Persson et al., 2021). Future studies could consider alternative personalization strategies, introduce time dependent modeling, and combining HRV with other information to improve the performance of HRV based assessment.

4. Conclusions

HRV has the potential to be a valuable marker for detecting driver fatigue. However, substantial progress is still required before HRV-based driver fatigue detection can be deployed in real life driving. Reviewed articles show that reduced HR is associated with fatigued driving states. However, when it comes to other HRV measures, the direction of change is not consistent. We believe the inconsistency could be introduced by the differences in causal factors and reference measurement for fatigue that were implemented in different studies. There is a need for more concrete knowledge about how HRV changes with different levels and causes of fatigue and their relation to driver performance. The performance of HRV based fatigue detection systems show a wide range of accuracy, the results are difficult to compare across all studies due to differences in the experiment setups. Reduced detection performance can be found in studies with large on-road experiments and subject-independent modeling. Using alternative personalization strategies, time dependent modeling, and utilizing other types of information could potentially contribute to more accurate detection in the future. Current findings from simulator and controlled on-road studies need to be further validated with real life driving studies.

CRediT authorship contribution statement

Ke Lu: Conceptualization, Methodology, Formal analysis, Investigation, Visualization, Writing – original draft, Writing – review & editing. **Anna Sjörs Dahlman:** Conceptualization, Writing – original draft, Writing – review & editing. **Johan Karlsson:** Conceptualization, Writing – original draft, Writing – review & editing, Project administration. **Stefan Candefjord:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This study was funded by a grant from the strategic vehicle research and innovation (FFI) program at Sweden's Innovation Agency (VINNOVA), grant number 2020-05157.

References

Abtahi, F., Anund, A., Fors, C., Seoane, F., Lindecrantz, K., 2018. Association of Drivers' sleepiness with heart rate variability: A Pilot Study with Drivers on Real Roads. In: Eskola, H., Väisänen, O., Viik, J., Hyttinen, J. (Eds.), EMBEC & NBC 2017, IFMBE Proceedings, 65. Springer Singapore, Singapore, pp. 149–152. https://doi.org/10.1007/978-981-10-5122-7_38.

Ahlstrom, C., Fors, C., Anund, A., Hallvig, D., 2015. Video-based observer rated sleepiness versus self-reported subjective sleepiness in real road driving. European Transport Research Review 7 (4), 38. https://doi.org/10.1007/s12544-015-0188-y.

Ahlström, C., Zemblys, R., Jansson, H., Forsberg, C., Karlsson, J., Anund, A., 2021. Effects of partially automated driving on the development of driver sleepiness. Accident Analysis & Prevention 153, 106058. https://doi.org/10.1016/j.aap.2021.106058.

Åkerstedt, T., Anund, A., Axelsson, J., Kecklund, G., 2014. Subjective sleepiness is a sensitive indicator of insufficient sleep and impaired waking function. Journal of Sleep Research 23 (3), 242–254. https://doi.org/10.1111/jsr.12158.

Appelhans, B.M., Luecken, L.J., 2006. Heart Rate Variability as an Index of Regulated Emotional Responding. Review of General Psychology 10 (3), 229–240. https://doi.org/10.1037/1089-2680.10.3.229.

Awais, M., Badruddin, N., Drieberg, M., 2014. In: a Non-invasive approach to Detect Drowsiness in a Monotonous Driving Environment, In: TENCON. https://doi.org/10.1109/TENCON.2014.7022356.

Billman, G.E., 2013. The LF/HF ratio does not accurately measure cardiac sympatho-vagal balance. Frontiers in Physiology 4 FEB, 26. https://doi.org/10.3389/fphys.2013.00026.

Boos, C.J., Vincent, E., Mellor, A., O'Hara, J., Newman, C., Cruttenden, R., Scott, P., Cooke, M., Matu, J., Woods, D.R., 2017. The Effect of Sex on Heart Rate Variability at High Altitude. Medicine and Science in Sports and Exercise 49 (12), 2562–2569. https://doi.org/10.1249/MSS.0000000000001384.

Buendia, R., Forcolin, F., Karlsson, J., Arne Sjöqvist, B., Anund, A., Candefjord, S., 2019. Deriving heart rate variability indices from cardiac monitoring—An indicator of driver sleepiness. Traffic Injury Prevention 20 (3), 249–254. https://doi.org/10.1080/15389588.2018.1548766.

Byeon, M.K., Han, S.W., Min, H.K., Wo, Y.S., Park, Y.B., Huh, W., 2006. A study of HRV analysis to detect drowsiness states of drivers. In: In: Proceedings of the Fourth IASTED International Conference on Biomedical Engineering, pp. 153–155.

Castaldo, R., Melillo, P., Bracale, U., Caserta, M., Triassi, M., Pecchia, L., 2015. Acute mental stress assessment via short term HRV analysis in healthy adults: A systematic review with meta-analysis. Biomedical Signal Processing and Control 18, 370–377. https://doi.org/10.1016/j.bspc.2015.02.012.

Chowdhury, A., Shankaran, R., Kavakli, M., Haque, M.M., 2018. Sensor Applications and Physiological Features in Drivers' Drowsiness Detection: A Review. IEEE Sensors Journal 18 (8), 3055–3067. https://doi.org/10.1109/JSEN.2018.2807245.

Chua, E.C.P., Tan, W.Q., Yeo, S.C., Lau, P., Lee, I., Mien, I.H., Puvanendran, K., Gooley, J. J., 2012. Heart rate variability can be used to estimate sleepiness-related decrements in psychomotor vigilance during total sleep deprivation. Sleep 35 (3), 325–334. https://doi.org/10.5665/sleep.1688.

European Parliament, Council of the European Union, 2019. Regulation (EU) 2019/2144 of the European Parliament and of the Council. https://eur-lex.europa.eu/eli/reg/2019/2144.

Egelund, N., 1982. Spectral analysis of heart rate variabilityas an indicator of driver fatigue. Ergonomics 25 (7), 663–672. https://doi.org/10.1080/00140138208925026.

Euro NCAP, 2017. Euro NCAP 2025 Roadmap. https://preview.thenewsmarket.com/Previews/NCAP/DocumentAssets/484064.pdf.

Franken, P., Dijk, D.-J., 2009. Circadian clock genes and sleep homeostasis. European Journal of Neuroscience 29, 9. https://doi.org/10.1111/j.1460-9568.2009.06723.x.

Fujiwara, K., Abe, E., Kamata, K., Nakayama, C., Suzuki, Y., Yamakawa, T., Hiraoka, T., Kano, M., Sumi, Y., Masuda, F., Matsuo, M., Kadotani, H., 2019. Heart Rate Variability-Based Driver Drowsiness Detection and Its Validation With EEG. IEEE

Transactions on Biomedical Engineering 66 (6), 1769–1778. https://doi.org/10.1109/TBME.2018.2879346.
Gonçalves, J., Bengler, K., 2015. Driver State Monitoring Systems– Transferable Knowledge Manual Driving to HAD. Procedia Manufacturing 3, 3011–3016. https://doi.org/10.1016/j.promfg.2015.07.845.
Hallvig, D., Anund, A., Fors, C., Kecklund, G., Åkerstedt, T., 2014. Real driving at night - Predicting lane departures from physiological and subjective sleepiness. Biological Psychology 101 (1), 18–23. https://doi.org/10.1016/j.biopsycho.2014.07.001.
Henelius, A., Sallinen, M., Huotilainen, M., Müller, K., Virkkala, J., Puolamäki, K., 2014. Heart rate variability for evaluating vigilant attention in partial chronic sleep restriction. Sleep 37 (7), 1257–1267. https://doi.org/10.5665/sleep.3850.
Hidalgo-Muñoz, A.R., Mouratille, D., Matton, N., Causse, M., Rouillard, Y., El-Yagoubi, R., 2018. Cardiovascular correlates of emotional state, cognitive workload and time-on-task effect during a realistic flight simulation. International Journal of Psychophysiology 128, 62–69. https://doi.org/10.1016/j.ijpsycho.2018.04.002.
Jung, S.J., Shin, H.S., Chung, W.Y., 2014. Driver fatigue and drowsiness monitoring system with embedded electrocardiogram sensor on steering wheel. IET Intelligent Transport Systems 8 (1), 43–50. https://doi.org/10.1049/iet-its.2012.0032.
Kaida, K., Takahashi, M., Åkerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., Fukasawa, K., 2006. Validation of the Karolinska sleepiness scale against performance and EEG variables. Clinical Neurophysiology 117 (7), 1574–1581. https://doi.org/10.1016/j.clinph.2006.03.011.
Kaida, K., Åkerstedt, T., Kecklund, G., Nilsson, J.P., Axelsson, J., 2007. Use of subjective and physiological indicators of sleepiness to predict performance during a vigilance task. Industrial Health 45 (4), 520–526. https://doi.org/10.2486/indhealth.45.520.
Khamis, N.K., Ismail, F.R., Hesse, B., Schramm, D., Maas, N., Koppers, M., Nuawi, M.Z., Deros, B.M., 2016. Suitability of heart rate recording as physiological measures tool to determine drivers' performance impairment: A preliminary study. Jurnal Teknologi 78 (6–9), 25–30. https://doi.org/10.11113/jt.v78.9143.
Koenig, J., Thayer, J.F., 2016. Sex differences in healthy human heart rate variability: A meta-analysis. Neuroscience and Biobehavioral Reviews 64, 288–310. https://doi.org/10.1016/j.neubiorev.2016.03.007.
Körber, M., Cingel, A., Zimmermann, M., Bengler, K., 2015. Vigilance Decrement and Passive Fatigue Caused by Monotony in Automated Driving. Procedia Manufacturing 3, 2403–2409. https://doi.org/10.1016/j.promfg.2015.07.499.
Kundinger, T., Sofra, N., Riener, A., 2020a. Assessment of the potential of wrist-worn wearable sensors for driver drowsiness detection. Sensors (Switzerland) 20, 4. https://doi.org/10.3390/s20041029.
Kundinger, T., Yalavarthi, P.K., Riener, A., Wintersberger, P., Schartmüller, C., 2020b. Feasibility of smart wearables for driver drowsiness detection and its potential among different age groups. International Journal of Pervasive Computing and Communications 16 (1), 1–23. https://doi.org/10.1108/IJPCC-03-2019-0017.
Lee, B.G., Lee, B.L., Chung, W.Y., 2015. Smartwatch-based driver alertness monitoring with wearable motion and physiological sensor, in. In: Proceedings of the Annual international Conference of the IEEE Engineering in Medicine and Biology Society, pp. 6126–6129. https://doi.org/10.1109/EMBC.2015.7319790.
Lee, H., Lee, J., Shin, M., 2019. Using wearable ECG/PPG sensors for driver drowsiness detection based on distinguishable pattern of recurrence plots. Electronics (Switzerland) 8, 2. https://doi.org/10.3390/electronics8020192.
Lenis, G., Reichensperger, P., Sommer, D., Heinze, C., Golz, M., Dössel, O., 2016. Detection of microsleep events in a car driving simulation study using electrocardiographic features. Current Directions in Biomedical Engineering 2 (1), 283–287. https://doi.org/10.1515/cdbme-2016-0063.
Leonhardt, S., Leicht, L., Teichmann, D., 2018. Unobtrusive vital sign monitoring in automotive environments—A review. Sensors (Switzerland) 18 (9), 1–38. https://doi.org/10.3390/s18093080.
Li, G., Chung, W.Y., 2013. Detection of driver drowsiness using wavelet analysis of heart rate variability and a support vector machine classifier. Sensors (Switzerland) 13 (12), 16494–16511. https://doi.org/10.3390/s131216494.
Lohani, M., Payne, B.R., Strayer, D.L., 2019. A review of psychophysiological measures to assess cognitive states in real-world driving. Frontiers in Human Neuroscience. https://doi.org/10.3389/fnhum.2019.00057.
Luque-Casado, A., Perales, J.C., Cárdenas, D., Sanabria, D., 2016. Heart rate variability and cognitive processing: The autonomic response to task demands. Biological Psychology 113, 83–90. https://doi.org/10.1016/j.biopsycho.2015.11.013.
Mahachandra, M., Yassierli, S., I.Z., Suryadi, K.,, 2012. Sensitivity of heart rate variability as indicator of driver sleepiness. In: In: 2012 Southeast Asian Network of Ergonomics Societies Conference (SEANES). IEEE. https://doi.org/10.1109/SEANES.2012.6299577.
Malik, M., John Camm, A., Thomas Bigger, J., Breithardt, G., Cerutti, S., Cohen, R.J., Coumel, P., Fallen, E.L., Kennedy, H.L., Kleiger, R.E., Lombardi, F., Malliani, A., Moss, A.J., Rottman, J.N., Schmidt, G., Schwartz, P.J., Singer, D.H., 1996. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. Circulation 93 (5), 1043–1065. https://doi.org/10.1161/01.cir.93.5.1043.
Matuz, A., van der Linden, D., Kisander, Z., Hernádi, I., Kázmér, K., Csathó, Á., Yan, H., 2021. Enhanced cardiac vagal tone in mental fatigue: Analysis of heart rate variability in Time-on-Task, recovery, and reactivity. PLOS ONE 16 (3), e0238670. https://doi.org/10.1371/journal.pone.0238670.
May, J.F., Baldwin, C.L., 2009. Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. Transportation Research Part F: Traffic Psychology and Behaviour 12 (3), 218–224. https://doi.org/10.1016/j.trf.2008.11.005.
Melo, H.M., Nascimento, L.M., Takase, E., 2017. Mental Fatigue and Heart Rate Variability (HRV): The Time-on-Task Effect. Psychology and Neuroscience 10 (4), 428–436. https://doi.org/10.1037/pne0000110.
Moak, J.P., Goldstein, D.S., Eldadah, B.A., Saleem, A., Holmes, C., Pechnik, S., Sharabi, Y., 2007. Supine low-frequency power of heart rate variability reflects baroreflex function, not cardiac sympathetic innervation. Heart Rhythm 4 (12), 1523–1529. https://doi.org/10.1016/j.hrthm.2007.07.019.
Murugan, S., Selvaraj, J., Sahayadhas, A., 2020. Detection and analysis: driver state with electrocardiogram (ECG). Physical and Engineering Sciences in Medicine 43 (2), 525–537. https://doi.org/10.1007/s13246-020-00853-8.
Nunan, D., Sandercock, G.R.H., Brodie, D.A., 2010. A Quantitative Systematic Review of Normal Values for Short-Term Heart Rate Variability in Healthy Adults. Pacing and Clinical Electrophysiology 33, 11. https://doi.org/10.1111/j.1540-8159.2010.02841.x.
Ohyama, S., Nishiike, S., Watanabe, H., Matsuoka, K., Akizuki, H., Takeda, N., Harada, T., 2007. Autonomic responses during motion sickness induced by virtual reality. Auris Nasus Larynx 34 (3), 303–306. https://doi.org/10.1016/j.anl.2007.01.002.
Ouzzani, M., Hammady, H., Fedorowicz, Z., Elmagarmid, A., 2016. Rayyan-a web and mobile app for systematic reviews. Systematic Reviews 5 (1), 1–10. https://doi.org/10.1186/s13643-016-0384-4.
Patel, M., Lal, S.K.L., Kavanagh, D., Rossiter, P., 2011. Applying neural network analysis on heart rate variability data to assess driver fatigue. Expert Systems with Applications 38 (6), 7235–7242. https://doi.org/10.1016/j.eswa.2010.12.028.
Pendleton, D.M., Sakalik, M.L., Moore, M.L., Tomporowski, P.D., 2016. Mental engagement during cognitive and psychomotor tasks: Effects of task type, processing demands, and practice. International Journal of Psychophysiology 109, 124–131. https://doi.org/10.1016/j.ijpsycho.2016.08.012.
Persson, A., Jonasson, H., Fredriksson, I., Wiklund, U., Ahlstrom, C., 2021. Heart Rate Variability for Classification of Alert Versus Sleep Deprived Drivers in Real Road Driving Conditions. IEEE Transactions on Intelligent Transportation Systems 22 (6), 3316–3325. https://doi.org/10.1109/TITS.2020.2981941.
Philip, P., Åkerstedt, T., 2006. Transport and industrial safety, how are they affected by sleepiness and sleep restriction? Sleep Medicine Reviews 10 (5), 347–356. https://doi.org/10.1016/j.smrv.2006.04.002.
Piccirillo, G., Ogawa, M., Song, J., Chong, V.J., Joung, B., Han, S., Magrì, D., Chen, L.S., Lin, S.-F., Chen, P.-S., 2009. Power spectral analysis of heart rate variability and autonomic nervous system activity measured directly in healthy dogs and dogs with tachycardia-induced heart failure. Heart Rhythm 6 (4), 546–552. https://doi.org/10.1016/j.hrthm.2009.01.006.
Pinto, J.R., Cardoso, J.S., Lourenço, A., Carreiras, C., 2017. Towards a continuous biometric system based on ECG signals acquired on the steering wheel. Sensors (Switzerland) 17 (10), 2228. https://doi.org/10.3390/s17102228.
Rahim, H.A., Dalimi, A., Jaafar, H., 2015. Detecting drowsy driver using pulse sensor. Jurnal Teknologi 73 (3), 5–8. https://doi.org/10.11113/jt.v73.4238.
SAE, 2016. Automated Driving - Levels of Driving Automation Are Defined in New Sae International Standard J3016, SAE international.
Saeb, S., Lonini, L., Jayaraman, A., Mohr, D.C., Kording, K.P., 2017. The need to approximate the use-case in clinical machine learning. Gigascience 6, 5. https://doi.org/10.1093/gigascience/gix019.
Sahayadhas, A., Sundaraj, K., Murugappan, M., 2012. Detecting Driver Drowsiness Based on Sensors: A Review. Sensors 12, 12. https://doi.org/10.3390/s121216937.
Salvati, L., d'Amore, M., Fiorentino, A., Pellegrino, A., Sena, P., Villecco, F., 2021. On-road detection of driver fatigue and drowsiness during medium-distance journeys. Entropy 23 (2), 135. https://doi.org/10.3390/e23020135.
Schömig, N., Hargutt, V., Neukum, A., Petermann-Stock, I., Othersen, I., 2015. The Interaction Between Highly Automated Driving and the Development of Drowsiness. Procedia Manufacturing 3, 6652–6659. https://doi.org/10.1016/j.promfg.2015.11.005.
Shaffer, F., Ginsberg, J.P., 2017. An Overview of Heart Rate Variability Metrics and Norms. Frontiers in Public Health 5. https://doi.org/10.3389/fpubh.2017.00258.
Shinar, Z., Akselrod, S., Dagan, Y., Baharav, A., 2006. Autonomic changes during wake-sleep transition: A heart rate variability based approach. Autonomic Neuroscience: Basic and Clinical 130 (1–2), 17–27. https://doi.org/10.1016/j.autneu.2006.04.006.
Sikander, G., Anwar, S., 2019. Driver Fatigue Detection Systems: A Review. IEEE Transactions on Intelligent Transportation Systems 20 (6), 2339–2352. https://doi.org/10.1109/TITS.2018.2868499.
Silvani, A., Calandra-Buonaura, G., Dampney, R.A.L., Cortelli, P., 2016. Brain-heart interactions: Physiology and clinical implications. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 374 (2067), 20150181. https://doi.org/10.1098/rsta.2015.0181.
Sollers, J.J., Sanford, T.A., Nabors-Oberg, R., Anderson, C.A., Thayer, J.F., 2002. Examining changes in HRV in response to varying ambient temperature. IEEE Engineering in Medicine and Biology Magazine 21 (4), 30–34. https://doi.org/10.1109/memb.2002.1032636.
Sommer, D., Golz, M., 2010. Evaluation of PERCLOS based current fatigue monitoring technologies, in: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC'10. pp. 4456–4459. doi:10.1109/IEMBS.2010.5625960.
Vicente, J., Laguna, P., Bartra, A., Bailón, R., 2016. Drowsiness detection using heart rate variability. Medical and Biological Engineering and Computing 54 (6), 927–937. https://doi.org/10.1007/s11517-015-1448-7.
Vogelpohl, T., Kühn, M., Hummel, T., Vollrath, M., 2019. Asleep at the automated wheel—Sleepiness and fatigue during highly automated driving. Accident Analysis and Prevention 126, 70–84. https://doi.org/10.1016/j.aap.2018.03.013.
Voss, A., Heitmann, A., Schroeder, R., Peters, A., Perz, S., 2012. Short-term heart rate variability - Age dependence in healthy subjects. Physiological Measurement 33 (8), 1289–1311. https://doi.org/10.1088/0967-3334/33/8/1289.

Voss, B.A., Schulz, S., Schroeder, R., Baumert, M., Caminal, P., 2009. Methods derived from nonlinear dynamics for analysing heart rate variability. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 367 (1887), 277–296. https://doi.org/10.1098/rsta.2008.0232.

Wang, L., Li, J., Wang, Y., 2019. Modeling and Recognition of Driving Fatigue State Based on R-R Intervals of ECG Data. IEEE Access 7, 175584–175593. https://doi.org/10.1109/ACCESS.2019.2956652.

Watling, C.N., Mahmudul Hasan, M., Larue, G.S., 2021. Sensitivity and specificity of the driver sleepiness detection methods using physiological signals: A systematic review. Accident Analysis and Prevention 150, 105900. https://doi.org/10.1016/j.aap.2020.105900.

Weinbeer, V., Frey, A., Feldhütter, A., Jarosch, O., Marberger, C., Radlmayr, J., 2018. Drowsiness and Fatigue in Automated Driving - Empirical Data for an Integrative Framework. Proceedings of the Human Factors and Ergonomics Society Europe.

Williamson, A., Lombardi, D.A., Folkard, S., Stutts, J., Courtney, T.K., Connor, J.L., 2011. The link between fatigue and safety. Accident Analysis and Prevention 43 (2), 498–515. https://doi.org/10.1016/j.aap.2009.11.011.

Wörle, J., Metz, B., Thiele, C., Weller, G., 2019. Detecting sleep in drivers during highly automated driving: The potential of physiological parameters. IET Intelligent Transport Systems 13 (8), 1241–1248. https://doi.org/10.1049/iet-its.2018.5529.

Zeng, C., Wang, W., Chen, C., Zhang, C., Cheng, B., 2020. Sex differences in time-domain and frequency-domain heart rate variability measures of fatigued drivers. International Journal of Environmental Research and Public Health 17 (22), 1–17. https://doi.org/10.3390/ijerph17228499.

Zheng, Y.L., Ding, X.R., Poon, C.C.Y., Lo, B.P.L., Zhang, H., Zhou, X.L., Yang, G.Z., Zhao, N., Zhang, Y.T., 2014. Unobtrusive sensing and wearable devices for health informatics. IEEE Transactions on Biomedical Engineering 61 (5), 1538–1554. https://doi.org/10.1109/TBME.2014.2309951.

Zwahlen, D., Jackowski, C., Pfäffli, M., 2016. Sleepiness, driving, and motor vehicle accidents: A questionnaire-based survey. Journal of Forensic and Legal Medicine 44, 183–187. https://doi.org/10.1016/j.jflm.2016.10.014.