applied sciences

MDPI

Article

Validation of Commercial Wearables for Driver Monitoring: Assessing Heart Rate Accuracy and Reliability for Early Detection of Sudden Illnesses

Hristo Radev [1,2,*] and Galidiya Petrova [1,2]

[1] Department of Electronics, Faculty of Electronics and Automation, Technical University of Sofia, Plovdiv Branch, 25 "Tsanko Dyustabanov" Str., 4000 Plovdiv, Bulgaria; gip@tu-plovdiv.bg
[2] Center of Competence "Smart Mechatronic, Eco- and Energy-Saving Systems and Technologies", 25 "Tsanko Dyustabanov" Str., 4000 Plovdiv, Bulgaria
* Correspondence: h_radev@tu-plovdiv.bg

Abstract

Sudden medical illnesses (SMI), including myocardial infarctions and occult arrhythmias, represent a primary etiology of catastrophic motor vehicle collisions. Although consumer wearable technologies provide a non-invasive modality for continuous driver monitoring, their measurement accuracy within dynamic automotive environments requires systematic evaluation. This investigation establishes an empirical validation paradigm for assessing the pulse-rate (PR) accuracy of consumer-grade wrist and finger photoplethysmography (PPG) devices against a reference heart rate (HR) value derived from an electrocardiographic (ECG) signal. To secure an uncompromised physiological reference, an analytical signal-processing pipeline was implemented, incorporating an Adaptive Layer-Dependent Thresholding Function (ALDTF) coupled with Maximal Overlap Discrete Wavelet Transform (MODWT) decomposition to suppress high-frequency motion artifacts and baseline wander (BW) in the ECG signal without inducing morphological distortion. Nineteen real-world driving sessions, totaling over 18 h of continuous physiological telemetry, were evaluated via Lin's Concordance Correlation Coefficient and Bland–Altman limit-of-agreement statistics. The findings demonstrate a distinct performance bifurcation governed by device form factor and internal data filtering. The evaluated smartwatch exhibited reliable tracking (MAE = 2.73 BPM) and preserved underlying autonomic variability, attaining distributional alignment with the ECG reference in 47% of trials. Conversely, ring-based form factors demonstrated substantial perfusion decoupling induced by isometric manual gripping forces, alongside aggressive proprietary smoothing filters that suppressed micro-fluctuations in heart rate variability (0% distributional alignment). These outcomes indicate that while wrist-worn form factors possess sufficient fidelity for detecting subclinical autonomic shifts, their implementation in automotive safety systems requires access to raw data to circumvent proprietary algorithmic filtering.

Keywords: driver monitoring systems; wearable devices; electrocardiography (ECG); photoplethysmography (PPG); heart rate variability (HRV); sudden medical illnesses (SMI); artifact suppression

Check for updates

Academic Editor: Emi Yuda

Received: 1 September 2026
Revised: 5 October 2026
Accepted: 10 October 2026
Published: 14 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

1.1. The Context: The Importance of Driver Physiological Monitoring for the Prevention of Sudden Cardiovascular and Neurological Incidents

Conventional road safety and accident prevention strategies have primarily targeted behavioral variables, notably fatigue, distraction, and aggressive driving [1]. Yet, a substantial and frequently fatal proportion of collisions stems from sudden driver incapacitation secondary to acute medical emergencies [2]. Unlike drowsiness, which manifests progressively, catastrophic physiological events strike without overt behavioral warning. Such critical incidents include myocardial infarction, stroke, or severe arrhythmias. These abrupt crises precipitate immediate incapacitation, leaving the vehicle entirely uncontrolled.

Continuous real-time physiological monitoring, specifically the analysis of heart rate (HR) and heart rate variability (HRV), offers a viable mechanism to detect the subclinical autonomic shifts that generally precede these acute conditions [3]. Cardiovascular and cerebrovascular emergencies constitute the primary etiologies of sudden driver incapacitation [4]. Predicting these events requires precise characterization of their prodromal physiological signatures—defined as the critical phase spanning between the onset of initial symptoms and the full manifestation of an acute medical emergency.

1.1.1. Acute Myocardial Infarction

This medical event typically presents as severe myocardial ischemia secondary to coronary obstruction. From a predictive standpoint, the autonomic nervous system responds to impending ischemia through pronounced sympathetic surges minutes to hours prior to the clinical event. This sympathetic hyperactivation is quantifiable as an abrupt decline in time-domain HRV metrics, notably the Standard Deviation of Normal-to-Normal intervals (SDNN, reflecting autonomic variability and total regulatory capacity) and the Root Mean Square of Successive Differences (RMSSD, capturing short-term beat-to-beat fluctuations driven predominantly by parasympathetic vagal tone). Consequently, a precipitous drop in RMSSD and SDNN serves as a primary subclinical indicator of autonomic collapse during nascent myocardial infarction. This critical decline is frequently accompanied by resting tachycardia [5].

1.1.2. Cardiac Arrhythmias

Conduction abnormalities pose a profound risk to drivers due to their potential to induce instantaneous syncope. Specifically, these critical conditions include atrial fibrillation (AFib), ventricular tachycardia, and ventricular fibrillation. The prodromal phases of these arrhythmias frequently feature ectopic activity, including premature ventricular or atrial contractions (PVCs/PACs). Granular temporal analysis of RR-intervals, utilizing methods such as Poincaré geometry, enables the isolation of these anomalies. Furthermore, the emergence of highly erratic, aperiodic RR-interval dynamics serves as a robust diagnostic marker for impending or occult AFib [6].

1.1.3. Cerebrovascular Incidents (Stroke)

Given that AFib is a primary catalyst for ischemic stroke, the continuous HR/HRV monitoring of silent AFib offers a direct proxy for evaluating stroke risk in real time [7]. Additionally, the acute hemodynamic fluctuations preceding cerebrovascular events often trigger distinct modulations in the baroreflex, phenomena that can be captured spectrally within the low-frequency (LF) components of the HRV [8,9].

1.1.4. Syncope (Sudden Loss of Consciousness)

Syncope, most commonly of vasovagal origin, is driven by transient cerebral hypoperfusion. Physiologically, its onset is heralded by acute bradycardia cascaded with sudden hypotension. High-resolution continuous monitoring is therefore capable of identifying this precipitous deceleration in pulse rate (PR) moments before full loss of consciousness occurs [10].

1.1.5. Vulnerable Risk Groups

Epidemiological evidence identifies specific demographics where continuous physiological monitoring yields the highest preventive value. Professional operators of heavy commercial vehicles face compounding occupational hazards, including prolonged sedentary exposure, chronic fatigue, and psychosocial stress. Consequently, this cohort exhibits elevated rates of obesity, hypertension, and obstructive sleep apnea—established precursors to acute cardiovascular and cerebrovascular events [11].

A secondary high-risk demographic comprises drivers over the age of 60. Aging induces a natural physiological attenuation in resting HRV and correlates with an exponentially higher prevalence of cardiovascular pathologies, notably occult arrhythmias like AFib [12]. Furthermore, targeted monitoring remains essential for individuals with established comorbidities, specifically ischemic heart disease, uncontrolled hypertension, and type 2 diabetes—the latter frequently complicated by autonomic neuropathy and silent myocardial infarctions [13]. Given these factors, the initial participant recruitment strategy for the current study aimed to prioritize an older driving cohort to maximize the probability of capturing subclinical markers of Sudden Medical Illnesses (SMI). However, due to practical challenges of recruiting elderly volunteers for prolonged multi-sensor driving experiments, the final studied cohort represents a predominantly middle-aged demographic.

1.1.6. Detection and Prediction via HR and HRV Analysis

Continuous physiological monitoring shifts vehicle safety paradigms from reactive mitigation to predictive diagnostics. Because the autonomic nervous system dynamically reflects the prodromal stages of physiological decompensation, the extraction of RR-interval metrics across multiple analytical domains yields robust predictive biomarkers [14].

Within the time domain, a pronounced attenuation of RMSSD (reflecting parasympathetic withdrawal and a loss of vagal tone by quantifying the magnitude of beat-to-beat variations between consecutive heart intervals) coupled with resting tachycardia typically indicates severe physiological stress or nascent ischemia. Concurrently, the incidence of ectopic activity can be quantified by abrupt interval variations exceeding 20% or 50 ms between consecutive heartbeats [15]. In the frequency domain, a dominant shift toward the low-frequency (LF) band and a marked elevation in the sympathovagal balance (the LF/HF ratio) frequently antedate hypertensive crises and acute coronary syndromes [16]. In the context of non-linear dynamics, a depression in signal complexity, measurable via Sample Entropy, signifies a critical loss of autonomic adaptability—a state strongly associated with impending sudden cardiac arrest [17].

1.1.7. Technological Solutions: Unobtrusive In-Cabin Monitoring and the Advantage of Wearables

Effective prediction of SMI requires continuous physiological monitoring that remains entirely non-intrusive to the driver. Conventional multi-lead clinical Holter systems are fundamentally impractical for driving operations. However, advances in microelectronics and digital signal processing now enable the integration of commercial wearable form factors [18]. Within this domain, ring-based devices offer specific anatomical advantages by providing continuous contact with the finger, potentially mitigating motion artifacts

during dynamic steering maneuvers. By acquiring high-resolution photoplethysmography (PPG) data and transmitting it for analytical processing, these devices generate continuous streams of RR-interval metrics [19]. Coupling such data acquisition with predictive classification architectures establishes a viable foundation for early-warning frameworks capable of triggering automated vehicle safety protocols upon detection of critical physiological precursors.

The widespread adoption of consumer wearables for in-cabin monitoring is propelled by distinct operational advantages. These devices specifically include smartwatches and smart rings. First, these devices provide continuous, transparent monitoring by eliminating cumbersome thoracic belts and clinical wiring, thereby sustaining high user adherence across extended driving intervals [16]. Second, consumer hardware exhibits high market penetration, providing an instantly accessible infrastructure without the specialized procurement barriers associated with research-grade instrumentation [18]. Furthermore, contemporary consumer platforms incorporate sufficient computational resources to execute lightweight edge analytics, facilitating localized alerts or autonomous vehicle interventions during critical anomalies. Finally, unlike proprietary medical monitoring systems, commercial wearables afford exceptional economic scalability, rendering their large-scale deployment feasible across commercial transport fleets and vulnerable demographic cohorts [19].

1.2. The Problem: Technological and Physiological Limitations of Commercial Wearable Devices in an Automotive Environment

Although consumer wearable devices present a pragmatic modality for continuous physiological monitoring, their measurement reliability within dynamic automotive environments remains insufficiently evaluated, particularly regarding the prospective detection of SMI. This technological bottleneck is multifaceted, arising from an intersection of mechanical interference, physiological constraints, algorithmic data distortion, and prevalent methodological gaps in the current research literature.

Foremost, the automotive cabin constitutes a challenging environment for optical PPG. Dynamic driving maneuvers introduce severe mechanical artifacts, caused primarily by isometric muscular tension on the steering wheel, fluctuating grip pressure, and high-frequency vibrations transmitted through the chassis.

Complex physiological barriers further complicate these mechanical hurdles. These challenges typically arise during acute systemic crises, including myocardial infarction, stroke, or severe arrhythmias. In such states, the body frequently experiences profound peripheral vasoconstriction, electromechanical decoupling, and rapid changes in Pulse Transit Time (PTT). This physiological response significantly attenuates the peripheral optical signal precisely when the critical pathological markers must be resolved.

Furthermore, to mitigate signal degradation during motion, commercial wearable devices predominantly rely on proprietary "black-box" filtering algorithms characterized by aggressive data smoothing and interpolation. While adequate for estimating mean HR during routine physical exercise, such heavy filtering suppresses the granular, high-frequency HRV metrics that serve as key diagnostic predictors for incipient arrhythmias or ischemia.

Consequently, the operational reliability of consumer devices in real-world driving scenarios is frequently overstated. Because validation protocols are predominantly executed under static or controlled conditions, systematic assessments evaluating Pulse Rate Variability (PRV) degradation during active vehicle control remain scarce. A comprehensive review of these biomechanical, physiological, and methodological constraints is detailed in Section 2 (Related Work).

Accordingly, rigorously establishing the efficacy (operational boundaries) of commercial wearables in automotive safety applications requires a high-precision analytical data processing. Consumer wearable data must be directly benchmarked against a reference ECG data acquired simultaneously in situ yet purified through transparent mathematical transformations that preserve the underlying signal morphology without relying on proprietary smoothing heuristics. Such an empirical framework is essential to delineate the precise operational thresholds beyond which consumer wearables cease to provide reliable data for vehicle safety systems.

1.3. Research Objectives

This investigation addresses the discrepancy between the theoretical utility of commercial wearable devices and their verified performance in dynamic automotive environments. Establishing whether mass-market smartwatches and smart rings can reliably detect SMI or progressive driver health deterioration requires a rigorous validation framework. Accordingly, this study pursues two primary, interconnected research objectives.

The first objective is focused on the development of a high-precision offline analytical pipeline designed to purify the reference ECG signal acquired under severe motion-noise conditions, because evaluating the PPG data against a clinical standard requires an uncompromised reference. However, the ECG signal is inherently contaminated by respiratory Baseline Wander (BW) and high-frequency electromyographic (EMG) interference generated by sustained steering maneuvers. To resolve this issue, the study implements a transparent ("white-box") offline signal-processing workflow utilizing non-causal filtering architectures, which prevent phase distortion. This pipeline integrates automated signal quality assessment (SQA) to isolate severe anomalies, combined with cascaded filtering designed to eliminate environmental noise while preserving R-peak morphology. Subsequent R-peak detection and artifact interpolation guided by Median Absolute Deviation (MAD) provide the millisecond-level accuracy required for reliable HRV extraction.

The second objective entails the quantitative evaluation of consumer wearables' accuracy using Lin's Concordance Correlation Coefficient (CCC) and Bland–Altman limit-of-agreement analysis. To overcome the limitations of conventional linear correlation metrics that fail to capture absolute deviations, this protocol applies rigorous biostatistical validation standards typical of medical instrumentation. Lin's CCC quantifies the absolute agreement between the consumer device data and the ECG reference value, penalizing systematic offsets. Concurrently, Bland–Altman analysis is utilized to identify systematic bias and establish the Limits of Agreement. This method reveals whether the device error increases during acute physiological shifts, such as stress responses or cardiac events. Furthermore, it determines whether the proprietary smoothing algorithms suppress the high-resolution variability essential for predictive diagnostics. Fulfillment of these objectives provides an empirical assessment regarding whether wrist and finger form factors possess sufficient technological maturity for integration into future vehicle safety systems.

2. Related Work

2.1. PRV vs. HRV During Motion

The proliferation of consumer wearables has facilitated the integration of PPG into continuous cardiovascular telemetry [20,21]. Such optical sensors generate PRV metrics, frequently employed as a surrogate for ECG-derived HRV [22,23]. Under controlled conditions and at rest, the current literature consistently confirms a high degree of agreement between PRV and HRV, rendering the two modalities practically interchangeable for basal physiological evaluation [22–26].

However, this equivalence deteriorates rapidly under kinetic conditions, with pronounced discrepancies emerging in complex psychomotor settings such as driving [27–30]. This divergence originates from the fundamental mechanics of data acquisition [31]: whereas ECG records instantaneous myocardial biopotentials, PPG measures the subsequent mechanical pulse wave propagating through the peripheral microvasculature [32]. During physical exertion or sympathetic arousal, this mechanical transmission is confounded by acute hemodynamic fluctuations, variable PTT, and sympathetically mediated peripheral vasoconstriction [33–36].

Within the vehicle cabin specifically, continuous steering adjustments, fluctuating grip pressure, and isometric muscular tension cause severe mechanical deformations upon the optical PPG waveform [27,37–40]. It is well established that the validity of PRV degrades substantially during dynamic motor activity [22]. However, the specific influence of driving-induced artifacts on the diagnostic reliability of consumer wearables remains inadequately addressed. This gap in the literature is particularly significant when considering their ability to isolate subclinical medical precursors [41].

2.2. Methodological Limitations in the Validation of Wearable Devices

The current literature evaluating consumer wearables often exhibits substantial methodological and statistical shortcomings [42–45], potentially leading to erroneous conclusions regarding measurement accuracy [26]. The main deficiency stems from experimental design [33,46]. Prevalent validation protocols assess wrist- and finger-worn devices predominantly under static resting conditions [47] or during rhythmic, predictable kinematics (e.g., treadmill running). Such paradigms fail to reproduce the aperiodic motion artifacts and complex mechanical disturbances endemic to real-world vehicle operation [42,48–50].

These environmental limitations are compounded by a pervasive statistical vulnerability: the over-reliance on the Pearson's correlation coefficient (r) as a primary metric for device accuracy [42,51,52]. Because Pearson's r quantifies merely the strength of linear covariance, it remains inherently blind to systematic bias and proportional mean offsets [43]. Consequently, a sensor exhibiting a constant measurement error (e.g., a continuous systemic offset of +10 BPM) may yield a falsely high correlation despite lacking clinical validity [42,43,52,53].

To robustly establish diagnostic utility, modern biostatistical standards require metrics that account for absolute agreement rather than mere covariance [43,53]. Therefore, rigorous validation requires the Lin's CCC along with Bland–Altman analysis to quantify systematic bias and establish limits of agreement between physiological extremes [42,43,47,54]. The lack of such rigorous, dynamically oriented validation frameworks in current research overestimates the technological readiness of consumer wearable devices for integration into critical, life-saving automotive applications [41,43,53].

2.3. The Necessity of Advanced Artifact Suppression in Reference Signals

Resolving the aforementioned methodological deficiencies and accurately quantifying the performance of wearable devices requires an uncompromised clinical ECG reference data. However, acquiring high-fidelity ECG signals during active driving operation remains technically demanding. The automotive environment inherently degrades signal integrity through two main sources. These include low-frequency BW [55] and high-frequency EMG interference generated by upper-limb motor engagement during steering [37]. Furthermore, the BW is primarily induced by postural dynamics and respiration [27,56].

Although modern ECG sensors (e.g., CardioBAN) incorporate Inertial Measurement Units (IMUs) to facilitate multi-sensor adaptive noise cancellation, kinematic-based filtering exhibits distinct limitations for high-resolution HRV extraction [27,57]. Vehicle arti-

facts do not primarily originate from gross torso accelerations, which IMUs capture with high effectiveness. Instead, they arise predominantly from localized skin-electrode micro-displacements and sustained isometric EMG interference [27,57,58]. Because kinematic sensors cannot capture these localized electromechanical transients, adaptive algorithms (e.g., Least Mean Squares) utilizing IMU data as a noise reference frequently fail to isolate artifacts comprehensively and tend to introduce phase distortions [27,57,58].

Similarly, conventional linear filtering techniques, such as standard infinite impulse response (IIR) or finite impulse response (FIR) implementations, have proven inadequate for this specific application. While effective in suppressing out-of-band spectral noise, these processing techniques inherently induce phase shifts and morphological changes in the QRS complex. Such structural distortions compromise the precise temporal localization of R-peaks, thereby degrading the millisecond-level precision required for reliable HRV quantification [55,59].

Consequently, establishing a robust reference signal without morphological degradation requires adaptive, non-linear techniques applied directly to the recorded raw ECG signals [60]. Offline analytical frameworks utilizing the Maximal Overlap Discrete Wavelet Transform (MODWT) coupled with adaptive layer-dependent thresholding functions (ALDTF) offer a highly effective alternative [60,61]. As non-causal and translation-invariant operators, these advanced wavelet methods circumvent phase shifting [61,62]. This methodology ensures the comprehensive suppression of dynamic environmental noise while preserving the temporal and structural integrity of the ECG signal, yielding an optimal reference for the subsequent benchmarking of consumer wearable devices [60,62].

3. Methodology

3.1. Experimental Setup and Protocol

An experimental setup was developed to acquire and synchronize multimodal physiological telemetry signals in real-world driving environments. The specific hardware topology, including the precise anatomical positioning of two research-grade sensors (CardioBAN and respiBAN) and the evaluated consumer wearables, has been detailed in the authors' previous paper [63].

This section outlines the core experimental parameters, cohort demographics, and algorithmic configurations applied in the present comparative analysis.

3.1.1. Participants

The study cohort consisted of 21 healthy volunteers. Individuals with acute or diagnosed medical conditions compromising driving ability were strictly excluded. In recruiting participants, priority was given to an older demographic to maximize the probability of capturing subclinical physiological markers associated with SMI. Written informed consent was obtained from all subjects prior to data collection. The experimental protocol was approved by the Ethics Committee for Research Involving Human Participants at the Technical University of Sofia (Protocol No. 4/20.10.2025).

3.1.2. Driving Protocol

Participants completed a standardized ~60-min driving session using their personal vehicles to maintain their natural driving behavior. The prescribed route incorporated diverse traffic conditions, passing through dense urban traffic in the city of Plovdiv, Bulgaria, high-speed ring roads, and highway sections. Synchronized multimodal telemetry was continuously recorded to provide comprehensive data on the driving environment. These data encompassed cabin temperature, humidity, GPS telemetry, and dual-perspective video from both driver-facing and road-ahead cameras. These recordings were essential for

cross-referencing physiological signal degradation with specific biomechanical artifacts, including steering maneuvers, braking events, and postural adjustments.

3.1.3. Instrumentation and Multi-Modal Sensors

The sensor array integrated reference research-grade sensors with consumer wearable devices [63] and is illustrated on Figure 1.

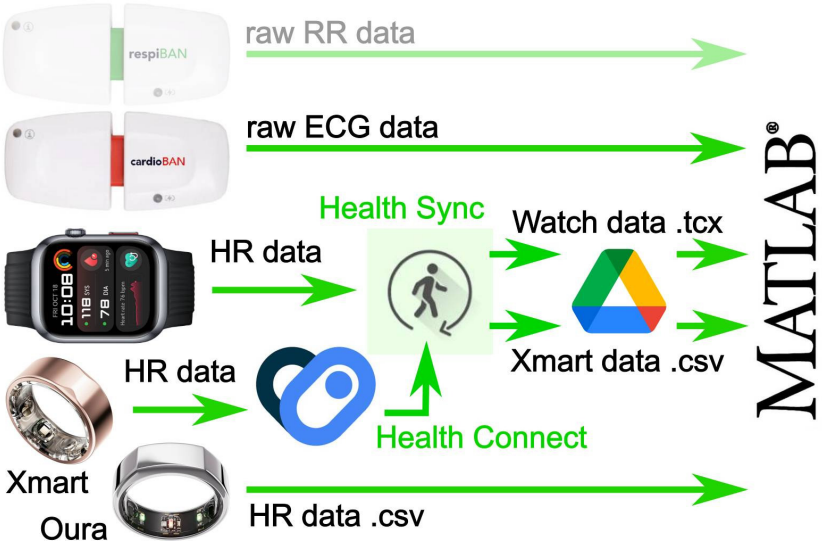


Figure 1. Experimental multi-modal sensor array and data synchronization architecture. The transparent respiBAN sensor is part of the experimental setup, but its respiratory data are not utilized in the current study. The arrows indicate the continuous flow of physiological data—from their initial generation by the wearable sensors, through intermediate cloud storage, to their final import into MATLAB as input data for synchronization and analytical processing.

The ECG reference signal was acquired using a CardioBAN BLE (PLUX Wireless Biosignals S.A., Lisbon, Portugal, hardware version: 1.0, firmware version: 1.7) configured in Lead I orientation and sampled at 1000 Hz. The respiratory signal was simultaneously recorded at 400 Hz using a respiBAN (PLUX Wireless Biosignals S.A., Lisbon, Portugal, hardware version: 1.0, firmware version: 1.6) inductive respiratory belt. It should be noted that these respiratory data fall outside the scope of the present analysis, as indicated by the transparent representation of the respiBAN sensor in Figure 1.

For consumer-grade telemetry, a Huawei Watch D2-8FA (Huawei Device Co., Ltd., Shenzhen, China, model: LCA-B10, firmware version: HarmonyOS 6.0.0.31) was affixed to the non-dominant wrist and two smart rings were attached to the fingers of the non-dominant hand, too. These included the Oura Ring Gen3 (Oura Health Oy, Oulu, Finland, firmware version: 3.4.3) and the Xmart smart ring (Xmart, Sofia, Bulgaria, firmware version: V0.7.0.7). This unilateral placement strategy was implemented to minimize high-frequency kinetic artifacts because the dominant upper limb typically executes dynamic secondary driving tasks (e.g., gear actuation, infotainment interaction, taking sips of beverage, instinctive body touching, and numerous other tasks), whereas the non-dominant limb maintains a relatively stable, isometric engagement with the steering wheel. Importantly, because these devices rely on closed software, future firmware updates by the manufacturers may change their internal filtering and, consequently, their accuracy.

3.2. Gold Standard Reference Processing and Signal Quality Assessment

To establish an uncompromised reference for evaluating consumer-grade wearables, the raw ECG signal acquired from the CardioBAN sensor was processed through an

automated signal quality assessment and artifact suppression pipeline, as depicted in Figure 2.

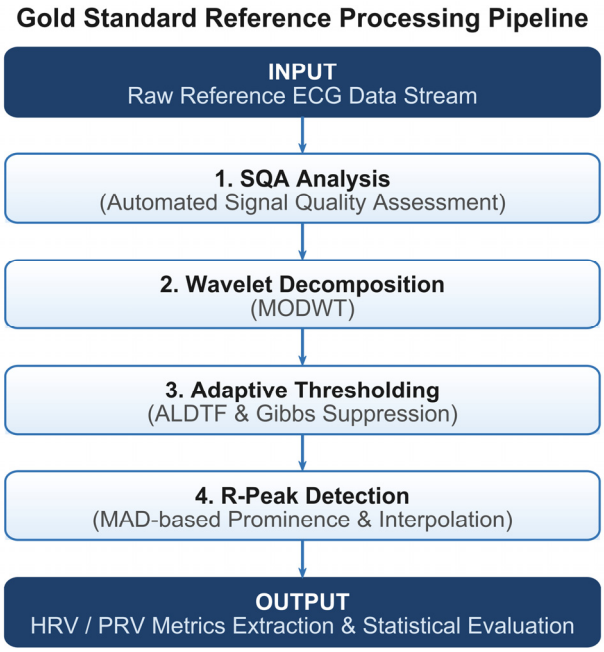


Figure 2. Schematic overview of the offline signal processing pipeline. The framework delineates the sequential stages from multimodal data acquisition, through mathematical artifact suppression (MODWT and ALDTF), to the extraction of HRV and PRV metrics for statistical benchmarking.

3.2.1. Automated Signal Quality Assessment (SQA)

The primary electrocardiographic reference (CardioBAN BLE) is inherently susceptible to motion artifacts, BW and EMG interference characteristic, of vehicle environments. Therefore, the raw ECG signal required dedicated pre-processing prior to metric extraction. The signal was first passed through an automated SQA module, beginning with DC offset removal using a 1-s moving median filter. The mean-centered signal was subsequently segmented into 1-s analysis windows to evaluate the local signal fidelity.

Individual epochs were classified as invalid and rejected if they met any of the following criteria: absolute amplitudes exceeding ± 3.0 mV (indicative of sensor saturation or detachment), minimal variance with a standard deviation (SD) < 0.01 (indicative of signal loss or flatlining), or excessive high-frequency noise (ringing) detected via an adaptive curve-length threshold. To mitigate edge-effect artifacts and ensure the integrity of adjacent data, the resulting binary rejection mask was extended both spatially and temporally, bridging minor valid gaps and imposing a 0.5-s exclusionary buffer around all corrupted segments. Full recordings demonstrating corruption rates above predefined tolerance thresholds were entirely excluded from subsequent comparative analyses to preclude instrumental bias.

3.2.2. Pre-Filtering Conditioning and Gibbs Suppression

Prior to wavelet decomposition, the masked signal undergoes pre-conditioning to mitigate spectral leakage and Gibbs-like boundary artifacts. The data array is initially zero-centered via global mean subtraction to eliminate the DC offset. To prevent boundary discontinuities, a 2-s Tukey (tapered cosine) window is applied, attenuating the signal extremities to zero, followed by symmetric zero-padding to isolate convolution-induced edge effects.

3.2.3. Baseline Wander Removal and Noise Suppression (MODWT & ALDTF)

Artifact suppression is performed by an ALDTF [60] applied within the MODWT domain. The MODWT framework was chosen for its translation (shift) invariance, avoiding phase distortion. The 'sym4' basis function was then applied because its near-symmetry closely matches the QRS complex, ensuring the precise timing of the electrocardiographic waveforms is maintained. This filtration strategy was implemented specifically to conserve critical ECG morphology—a fundamental prerequisite for isolating subclinical predictive markers of vehicle medical emergencies.

To eliminate BW, an analytical energy minimum is identified across decomposition levels [64]. This threshold, known as the 'Stop Point', is physiologically constrained between levels 9 and 11 (defaulting to level 10 analytically) for a 1000 Hz sampling rate. Approximation coefficients beyond this threshold are zeroed. To attenuate high-frequency kinematic and EMG artifacts across the lower decomposition levels (1 through 5, with default value of 3), the noise variance (σ) is robustly estimated utilizing the MAD:

$$\sigma = \frac{median\left(\left|d_j\right|\right)}{0.6745} \quad (1)$$

A universal threshold, defined as $\lambda = \sigma \times \sqrt{\ln N}$, is applied by an adaptive nonlinear transfer function incorporating a hyperbolic tangent. This formulation operates as soft thresholding in the lower decomposition layers to suppress pseudo-Gibbs oscillations, progressively transitioning to hard thresholding in the higher layers to maintain waveform amplitude. Finally, to prevent morphological over-smoothing, the shaping parameter α was empirically tuned to 1.5 and applied globally per session within the non-linear transfer function.

3.2.4. Robust R-Peak Detection and Reference Derivation

Upon inverse MODWT reconstruction and the restoration of physical millivolt units (by DC mean re-addition), cardiac cycles are localized through a peak-detection algorithm driven by MAD-derived prominence thresholds and physiological refractory constraints. A terminal SQA validation excludes any R-peaks situated within previously defined corruption masks. The verified R-R intervals are subsequently transformed into instantaneous heart rate (IHR) series. To establish the final reference values of HR, these series are subjected to artifact rejection (excluding values outside the physiological 40–200 BPM window), moving median filtration, and linear interpolation.

3.3. Multimodal Synchronization

While the hardware and integration framework were detailed previously [63], this full-scale phase introduces a critical update: transitioning the CardioBAN data from Excel to native MATLAB .mat structures. Due to variable traffic conditions, some driving sessions are extended up to 90 min, generating over 3 million rows of ECG data per session at a 1000 Hz sampling rate. Because Excel worksheets strictly limit capacity to 1,048,576 rows, traditional spreadsheets became unfeasible for these high-sampling resolution arrays. Consequently, the pipeline was upgraded to directly import high frequency .h5 streams into optimized .mat structures, ensuring computational efficiency and preventing memory bottlenecks.

3.3.1. Master Time Reference and Initial Alignment

The Huawei Watch D2 timestamps served as the master reference axis to ensure continuous temporal reference, given that smart rings were occasionally omitted due to anatomical sizing constraints. Consumer smart rings were aligned using device-specific

proximity thresholds designed to account for vendor logging intervals and cloud synchronization latencies. Specifically, the Oura Ring was synchronized using a nearest-neighbor temporal threshold of $\leq$2 s relative to the master watch timestamps, which suits its nominal 5-s logging interval in exercise mode while preventing data duplication. For the Xmart Ring, a slightly wider threshold of $\leq$3 s was applied to accommodate its more irregular reporting intervals, supplemented by automated data-bridging routines to seamlessly handle intermediate gaps. These tailored thresholds effectively compensate for minor Bluetooth and cloud export latencies inherent to each manufacturer's proprietary mobile ecosystem, ensuring precise temporal matching against the master reference.

This nearest-neighbor approach was chosen specifically to avoid resampling the data to a uniform 1 Hz grid. Interpolating the sparse outputs from commercial devices would introduce synthetic data points that were never actually recorded. Additionally, since their internal algorithms are proprietary, accurately replicating their averaging kernels is impossible. Therefore, the 2–3 s tolerance serves only to align the asynchronous timestamps without fabricating physiological data.

3.3.2. Integration and Boundary Clipping with the Reference Standard

A comparison MATLAB (The MathWorks Inc. MATLAB version: 24.1.0 (R2024a). Natick, MA, USA) script integrates high-resolution research-grade CardioBAN data (.mat structures containing instantaneous HR and datetime values) synchronized with the master timeline via nearest-neighbor interpolation based on timestamps. To eliminate boundary artifacts or desynchronization before and after driving, a strict boundary clipping protocol is enforced. Specifically, the script detects valid data boundaries defined by the start and stop indices of the CardioBAN reference recording, and all auxiliary data streams are subsequently clipped to match these exact reference start and end markers, ensuring comparison over 100% overlapping windows. Furthermore, any temporal points falling outside this active window or missing due to Bluetooth disconnection are systematically assigned NaN (Not a Number) values.

3.4. Statistical Evaluation Protocol

To rigorously quantify the accuracy, agreement, and distribution characteristics of the consumer-grade wearables against the research-grade CardioBAN reference standard, a comprehensive statistical protocol was established. The evaluation encompasses magnitude and percentage error metrics, including Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE), alongside concordance and agreement estimators such as Pearson's correlation coefficient (r), Lin's CCC, systematic bias, and Bland–Altman Limits of Agreement. Furthermore, the protocol incorporates physiological variability comparison through the assessment of PRV from wearables versus reference HRV by the standard deviation of HR (SD of HR), as well as non-parametric hypothesis testing utilizing Wilcoxon signed-rank tests for distributional divergence from the reference and pairwise comparisons among the evaluated wearables. While correlational coefficients evaluate proportional trend tracking, demonstrating the absence of statistically significant deviation strictly requires that the wearable device reproduces the identical statistical distribution and median of the reference ECG. This guarantees the absence of persistent systematic bias (e.g., algorithmic overestimation or underestimation) across thousands of continuous beat-to-beat measurements.

Regarding the statistical treatment of missing data, any temporal points assigned NaN values due to above mentioned reasons were excluded from the error metrics calculations via pairwise deletion. This approach ensures that quantitative accuracy and agreement estimators, such as MAE, RMSE, and Bland–Altman limits were derived exclusively from

concurrent, valid observations, thereby preventing imputation bias from artificially inflating the reported performance metrics.

4. Results

4.1. Participants Data

Twenty-one healthy volunteer drivers completed the 60-min real driving protocol. All participants were aged between 33 and 70 years. The study included 8 women, with average age of 46.0 years (SD = 11.55) and average body mass index (BMI) of 20.54 (SD = 3.52), and 13 men with average age of 47.31 years (SD = 12.77) and average BMI of 26.06 (SD = 5.06). Skin tones ranged from type one to three on the Fitzpatrick Skin Type scale. Our initial methodological aim was to focus exclusively on an older demographic. However, the 60-min instrumented driving protocol proved physically and psychologically demanding, which significantly limited volunteer availability. As a result, the final convenience sample naturally shifted toward a middle-aged distribution. Participants demographics and details are presented in Table 1.

Table 1. Participants characteristics.

Participants	Men (n = 13)		Women (n = 8)		Total (n = 21)	
	Mean (SD)	**Range**	**Mean (SD)**	**Range**	**Mean (SD)**	**Range**
Age (years)	47.31 (12.77)	33–70	46.00 (11.55)	36–68	46.81 (12.04)	33–70
Weight (kg)	84.31 (17.60)	64–115	60.25 (12.15)	41–80	75.14 (19.32)	41–115
Height (cm)	179.62 (6.11)	170–189	170.50 (7.45)	160–184	176.15 (7.98)	160–189
BMI (kg/m^2)	26.06 (5.06)	20.2–37.55	20.54 (3.52)	16.02–27	23.96 (5.22)	16.02–37.55

The successful deployment and physical compatibility of the measurement equipment were monitored across all N = 21 participants during the experimental protocol. The matrix (Figure 3) illustrates the distribution of device fittings relative to the individual anthropometric characteristics of the volunteers across four primary modality units: the reference electrocardiography sensor CardioBAN, the commercial smartwatch Huawei Watch, and the smart rings Oura Ring and Xmart Ring.

Device	Device Usage [%]	Participant #																				
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
CardioBAN	100	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Huawei Watch	100	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Oura Ring	85.7	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	NA	NA	Y	Y	Y	NA	Y	Y	Y
Xmart Ring	81.0	Y	Y	Y	NA	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	NA	NA	Y	Y	Y	NA	Y

Figure 3. Device usage matrix showing the successful deployment and sizing compatibility of wearable devices across all 21 participants. The symbol "#" denotes the participant identification number. Green cells marked with "Y" (Yes) indicate successful device deployment, while red cells marked with "NA" (Not Applicable) indicate unsuccessful deployment due to sizing incompatibility.

The reference ECG sensor CardioBAN and the commercial smartwatch Huawei Watch both achieved a maximum usability rate of 100% (21 out of 21 successful deployments, denoted by green "Y" cells). The universal design of the watch straps allowed for secure wrist fixation across all participants without encountering anatomical limitations. In contrast to wrist-worn devices, smart rings require a precise match with the user's finger dimensions. Due to individual anatomical differences and the absence of exact sizing options within the available test kits, a subset of participants could not be fitted with these devices (NA "red" values). This evaluation emphasizes that while wrist-worn form factors

and chest straps exhibit full universality, the integration of smart rings into large-scale studies is severely limited by sizing availability, as anatomical mismatches directly prevent hardware deployment.

4.2. Data Collection and Synchronization Statistics

For HR measurements, a maximum of 66,265 s of data (equivalent to 13,253 timestamps at 5-s intervals) could have been collected per device over the full test protocol. The CardioBAN reference standard recorded the complete possible volume of 13,253 HR observations. Among the tested commercial devices, the Huawei Watch also logged a total of 13,253 records, the Oura Ring recorded a total of 11,327 observations, and the Xmart Ring captured a total of 10,635 values. Collectively, these durations provided approximately 18 h and 24 min of valid, real-world data for the purposes of the study.

To ensure rigorous data integrity under real-world driving conditions, the data processing pipeline incorporates an automated algorithmic workflow that evaluates signal fidelity across all processing stages. Rather than relying solely on initial pre-filtration (SQA), the evaluation computes an Estimated Algorithmic Error Rate (%) at the conclusion of the processing pipeline, which quantifies the exact percentage of the signal deemed unusable or rejected. Those percentages are listed in Table 2 for each session.

Table 2. Quantitative evaluation of the automated Signal Quality Assessment and Algorithmic Error Rates across all 21 experimental driving sessions, extracted from individual session reports.

Session	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Data Loss [%]	1.18	0.72	1.79	0.57	7.96	3.27	24.93	3.29	0.23	2.25	0.93	0.87	4.07	0.71	3.76	0.49	0.03	1.99	1.1	0.36	0.67

To synthesize and evaluate the structural characteristics of the dataset generated by the automated algorithmic pipeline, the percentage error metrics from Table 2 were mapped onto a compact horizontal box-and-whisker distribution plot (Figure 4).

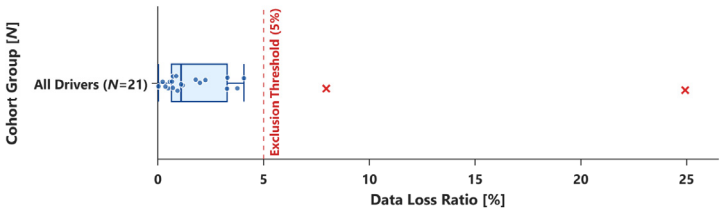


Figure 4. Horizontal box-and-whisker distribution plot illustrating the overall error distribution across the cohort (N = 21). Statistical outliers (representing the excluded sessions) are marked with a red '×' symbol.

The visualization of Figure 4 exhibits a pronounced right-skewed distribution across the cohort (N = 21). The central interquartile range, bounded by the 25th and 75th percentiles, demonstrates that for the vast majority of sessions, total algorithmic error remains exceptionally minimal (clustering well below 4%). Enforcing the predefined 5% Exclusion Threshold (denoted by the vertical dashed red line), sessions surpassing this global quality boundary were excluded from the cohort's results: session 5 and session 7. These sessions are explicitly identified on the distribution plot as red crosses. Error rates exceeding the 5% boundary do not stem from minor stochastic noise. Instead, these exclusions highlight the reality of severe data acquisition failures during genuine naturalistic driving. To keep the experimental conditions as authentic as possible, we intentionally avoided intervening in the participants' natural habits or driving styles. This approach exposed the recordings to the full unpredictability of real-world driving. A detailed review of the video telemetry

confirmed that the data loss in Sessions 5 and 7 was not caused by hardware or communication malfunctions, but rather by massive biomechanical interference from unconstrained physical movements. For example, in Session 5, an older participant (68 years old) made frequent postural adjustments due to the fatigue of prolonged driving, alongside expressive conversational gesturing. This participant also displayed a highly erratic driving style, making far more abrupt, high-effort steering corrections than the cohort average. Similarly, in Session 7, a participant with a higher BMI engaged in repetitive secondary tasks, such as reaching for their face. Because of the participant's physical build, the natural mobility of subcutaneous soft tissues during these large arm movements made the sensor highly unstable. In both scenarios, these unpredictable but entirely natural behaviors caused severe EMG noise and temporarily detached the ECG electrodes located just below the chest. This resulted in signal artifacts and amplitude saturation, which our automated SQA module correctly identified and rejected. Retaining these compromised recordings would inject instrumental noise into downstream statistical results.

4.3. Micro-Level Case Study: Detailed Analysis of Session 9

To illustrate the practical application of the proposed processing pipeline and algorithmic workflow under real-world driving conditions, Session 9 is presented here as a representative high-fidelity case study. Automated SQA performed on the raw ECG stream confirmed an exceptionally low data loss ratio of 0.23%, reflecting minimal motion artifact interference during the experimental route (Figure 5).

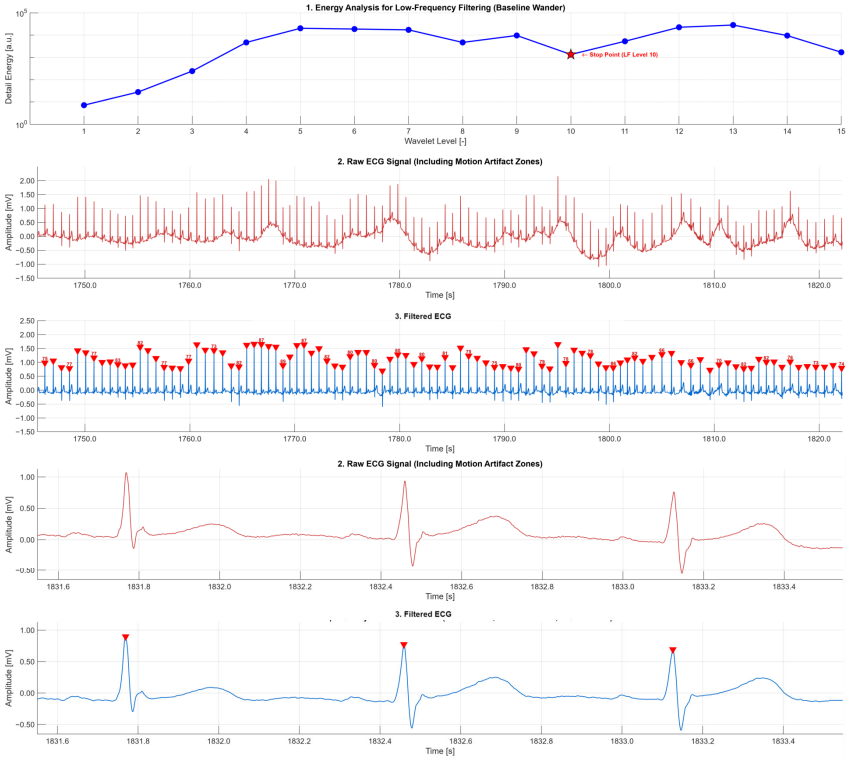


Figure 5. Micro-level signal processing and peak detection workflow for Session 9, generated by the automated algorithmic pipeline: (**top panel**) wavelet detail energy analysis up to level 15, identifying the optimal stop point at Level 10 for BW removal; (**upper-middle panel**) raw ECG signal containing baseline drift and motion artifact zones over a multi-second recording window; (**lower-middle panel**) filtered ECG signal with 0.00% SQA rejection rate alongside robust R-peak detection (red triangles) and IHR annotation; (**bottom two panels**) zoomed-in high-resolution views comparing the raw and filtered morphological waveforms and precise R-peak localization (84 BPM, 78 BPM, etc. marked by the red triangles and red number) at the individual heartbeat level. Blue lines represent the detail energy and filtered ECG, while red lines indicate the raw ECG signal.

Following adaptive BW removal via Stationary Wavelet Transform (Level 10) and high-frequency noise suppression, robust peak detection via MAD successfully identified all valid R-peaks without false triggers or missed beats. The automatically generated session report (9_ECG_HRV_Report.txt) documents a mean HR of 79.2 BPM, with stable global time-domain metrics including an overall variability (SDNN) of 75.3 ms and a parasympathetic tone (RMSSD) of 37.4 ms. Furthermore, non-linear Poincaré mapping and Welch power spectral density estimation yielded an SD1/SD2 ratio of 0.256 and an LF/HF ratio of 12.396, effectively capturing autonomic nervous system modulation and pronounced sympathetic dominance typical for active driving demands (Figure 6). Micro-variability tracking via 5-min rolling windows further demonstrated the pipeline's capability to continuously monitor driver autonomic responses, validating the recording's suitability for high-precision analysis.

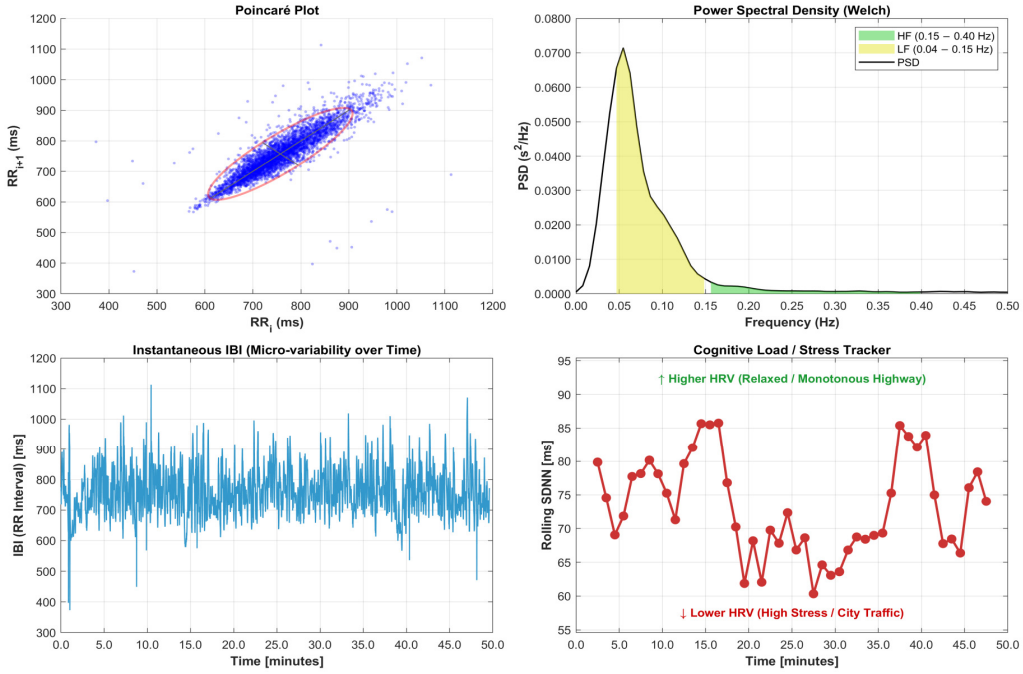


Figure 6. HRV, cognitive load, and autonomic stress analysis derived from Session 9: (**top-left**) Poincaré plot displaying short-term (SD1 = 26.5 ms) and long-term (SD2 = 103.2 ms) non-linear variability; (**top-right**) Power Spectral Density estimated via Welch's method, illustrating Low-Frequency (LF, 0.04–0.15 Hz) and High-Frequency (HF, 0.15–0.40 Hz) spectral power bands; (**bottom-left**) instantaneous inter-beat interval (IBI/RR) micro-variability over the ~50-min recording duration; (**bottom-right**) rolling 5-min SDNN cognitive load and stress tracker reflecting autonomic fluctuations corresponding to driving conditions (e.g., elevated HRV during relaxed highway driving versus reduced HRV under high-stress city traffic). Additionally, in the top-left panel, the red ellipse represents the standard deviation fitting of the data. In the top-right panel, the yellow and green shaded areas explicitly denote the LF and HF spectral bands. In the bottom-right panel, the red circles represent the individual SDNN values calculated for each rolling window.

To further evaluate the dynamic tracking fidelity and temporal alignment of the wearable monitors under real-world driving conditions, Figure 7 illustrates the continuous HR time-series comparison for Session 9.

The reference signal acquired by the CardioBAN ECG sensor (depicted in black) captures rapid physiological fluctuations and transient autonomic responses throughout the protocol. Among the tested consumer-grade devices, the Huawei Watch (red profile) exhibits superior tracking capabilities, closely mirroring the reference waveform across sharp peaks and drops with a low MAE = 2.60 BPM and strong concordance CCC = 0.766. In

contrast, the Oura Ring (blue profile) demonstrates adequate trend tracking but noticeable damping and underestimation during rapid HR elevations. The Xmart Ring (green stepped profile) displays significant quantization and temporal lag, reflecting the limitations of its raw PPG extraction during real-world driving conditions.

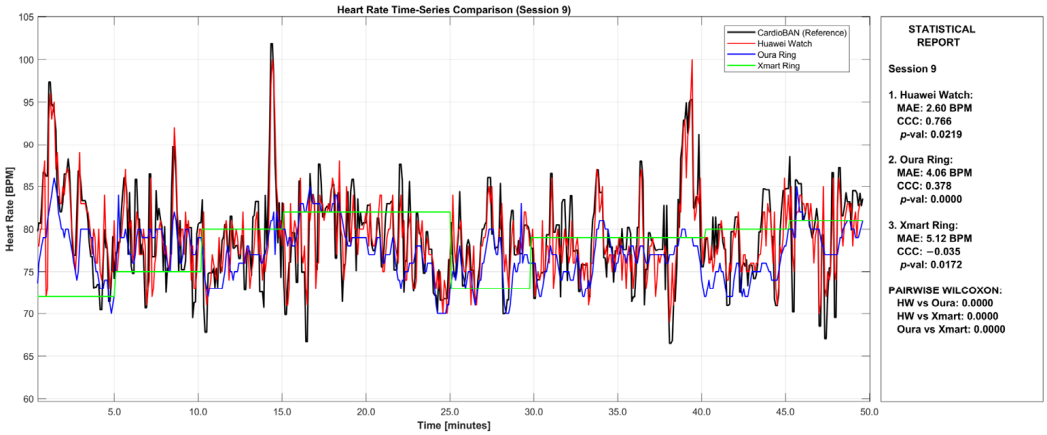


Figure 7. Comparative HR time-series and device synchronization plot for Session 9 over the ~50-min driving protocol: (main plot) continuous beat-to-beat and smoothed HR (BPM) profiles comparing the CardioBAN reference standard (black line) against the Huawei Watch (red line), Oura Ring (blue line), and Xmart Ring (green stepped line); (top-right popup) statistical export summary detailing individual device performance (MAE and Lin's CCC) alongside pairwise Wilcoxon test *p*-values.

To evaluate whether the HR distributions captured by the consumer-grade devices statistically deviated from the reference standard, non-parametric Wilcoxon signed-rank tests were performed for Session 9. The statistical evaluation revealed no statistically significant difference between the Huawei Watch and the CardioBAN reference ($p = 0.0219$), confirming high distributional alignment. In contrast, both the Oura Ring ($p = 0.0000$) and the Xmart Ring ($p = 0.0172$) exhibited statistically significant divergences from the reference. Furthermore, pairwise Wilcoxon comparisons between the wearables confirmed highly significant performance stratification (Huawei vs. Oura: $p = 0.0000$; Huawei vs. Xmart: $p = 0.0000$; Oura vs. Xmart: $p = 0.0000$), corroborating the visual disparities observed in the time-series profile.

To evaluate whether PR data from consumer wearables possess sufficient accuracy to capture subtle physiological fluctuations (such as stress- and fatigue-induced autonomic shifts) relative to the reference standard, dynamic HRV was quantified via the standard deviation of HR (SD of HR), as presented in Figure 8. This metric demonstrates the pipeline's capability to continuously monitor driver autonomic responses, validating the recording's suitability for high-precision analysis.

During Session 9, the CardioBAN reference standard recorded basal physiological variability of 5.55 BPM. Among the tested devices, the Huawei Watch successfully mirrored this dynamic range with an SD of HR of 4.60 BPM, demonstrating that wrist-PPG preserves core autonomic micro-fluctuations. In contrast, both the Oura Ring (SD of HR = 3.15 BPM) and the Xmart Ring (SD of HR = 3.46 BPM) exhibited greatly attenuated variability values. This significant suppression indicates that ring-based form factors apply aggressive smoothing algorithms that dampen transient physiological spikes, limiting their reliability for high-resolution stress and fatigue monitoring without prior calibration.

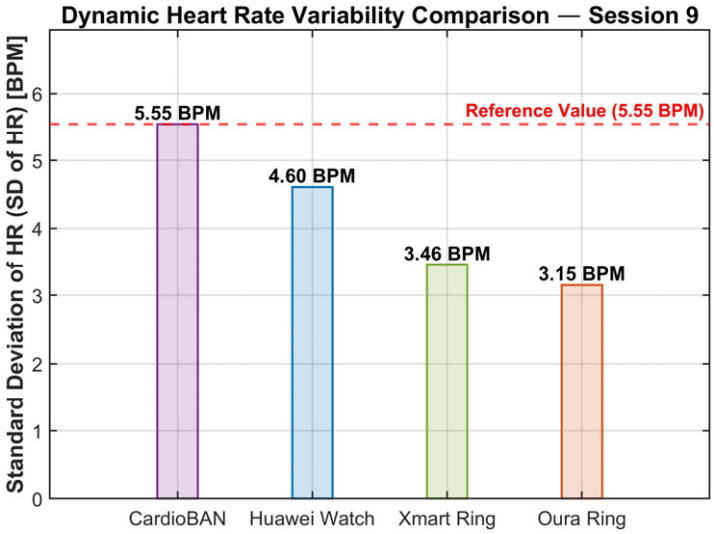


Figure 8. Comparison of dynamic HRV, quantified via the standard deviation of HR (SD of HR), between the CardioBAN reference standard and consumer wearable devices during Session 9. The red dashed line denotes the reference value (5.55 BPM), illustrating the attenuation of physiological micro-fluctuations in ring-form factor devices.

To substantiate the visual observations from the time-series comparison and to provide a comprehensive quantitative evaluation of device accuracy during the representative experimental run, key performance indicators for Session 9 are summarized in Table 3. The metrics illustrate the precise degree of concordance and statistical deviation between each consumer-grade wearable and the CardioBAN reference standard.

Table 3. Performance metrics for a representative experimental session (Session 9) comparing consumer wearable devices against the CardioBAN reference standard. The Wilcoxon signed-rank test *p*-value denotes the statistical significance of the measurement differences relative to the reference standard.

Device	MAE [BPM]	MAPE [%]	RMSE [BPM]	Bias [BPM]	Lin's CCC	*p*-Value (vs. Ref.)
Huawei Watch	2.60	3.31	3.49	−0.24	0.766	0.0219
Oura Ring	4.06	5.04	5.28	−2.03	0.378	<0.001
Xmart Ring	5.12	6.42	6.71	−0.88	−0.035	0.0172

Following the in-depth micro-level validation of an individual high-fidelity session, the macro-level evaluation was extended across the entire validated cohort (N = 19) to benchmark consumer wearables performance against the reference ECG sensor.

4.4. Macro Level: Overall Performance in Measuring HR

Following the rigorous data curation protocol and the systematic exclusion of compromised recordings (Sessions 5 and 7) due to excessive algorithmic error rates, the comparative performance of the tested consumer-grade wearable monitors against the reference ECG sensor (CardioBAN) was evaluated across the finalized cohort (N = 19 valid sessions). The agreement, error distribution, and concordance were comprehensively assessed using global Bland–Altman analyses presented in Figure 9, session-level box-and-whisker metrics visualized in Figure 10, and aggregated statistical parameters (Mean ± SD) in Table 4.

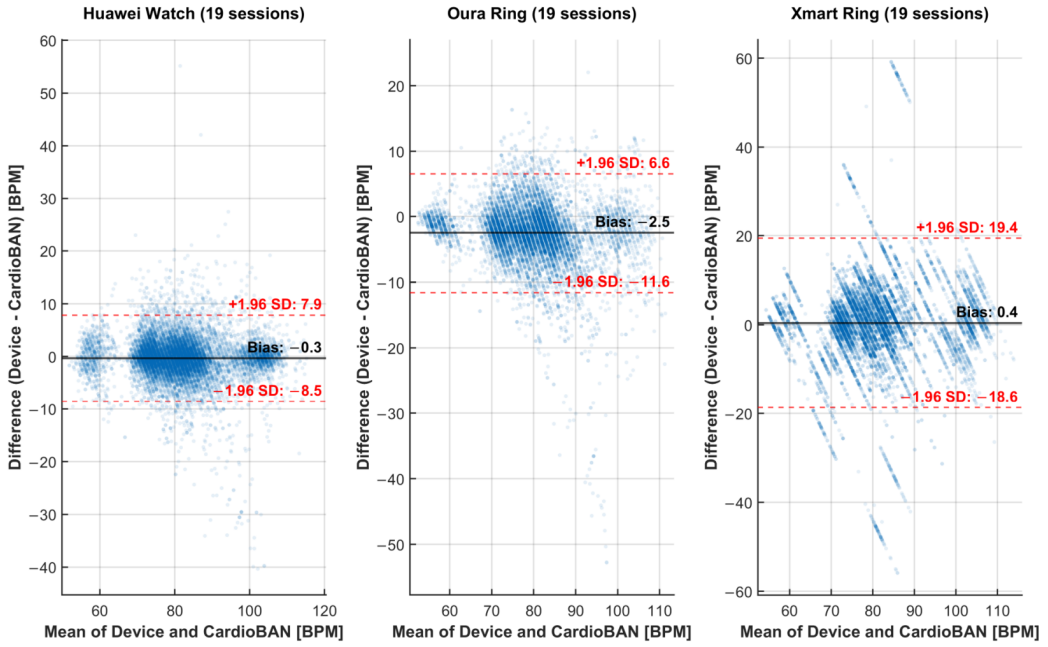


Figure 9. Global Bland–Altman agreement analysis comparing HR (BPM) measurements between the reference (CardioBAN) and three tested wearable devices—Huawei Watch, Oura Ring, and Xmart Ring—across the validated cohort (N = 19). The blue scattered dots represent the individual paired measurement differences. Solid black lines indicate the mean systematic bias, while red dashed lines and the corresponding red text represent the 95% limits of agreement (±1.96 SD).

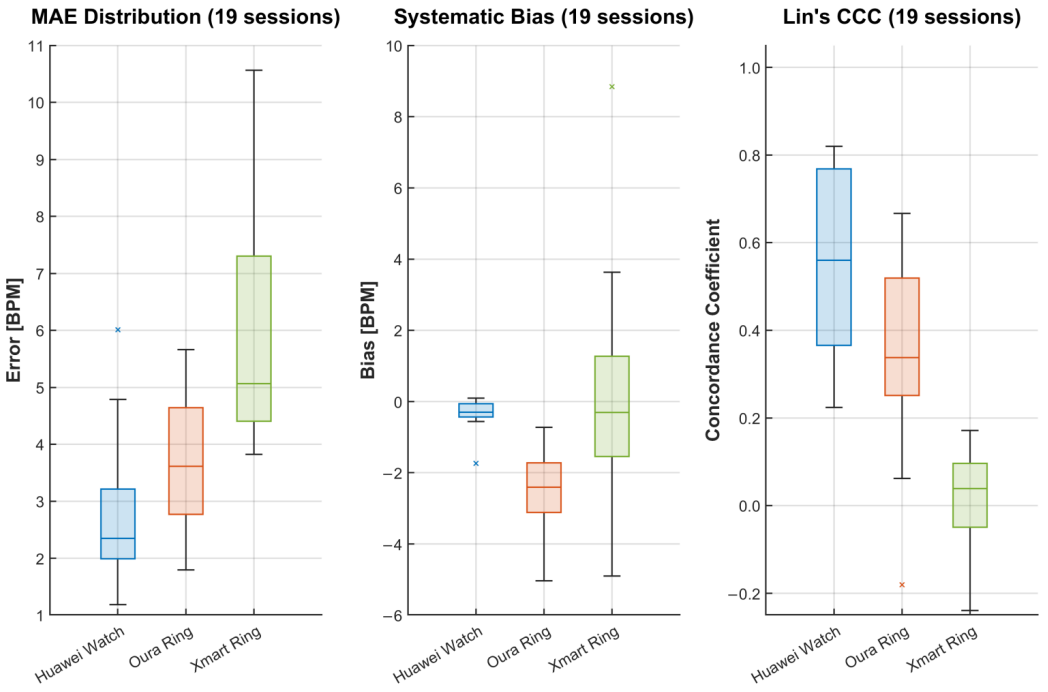


Figure 10. Distribution of session-level performance metrics across the 19 validated sessions for the Huawei Watch, Oura Ring, and Xmart Ring: (**left**) MAE, (**middle**) Systematic Bias, and (**right**) Lin's CCC. The '×' symbols indicate the statistical outliers for each respective device.

Table 4. Summary of global performance metrics (Mean ± SD) comparing the tested wearable devices against the CardioBAN reference across the validated cohort (N = 19 sessions).

Device	MAE [BPM]	MAPE [%]	RMSE [BPM]	Bias [BPM]	Lin's CCC
Huawei Watch	2.73 ± 1.19	3.37 ± 1.48	3.90 ± 1.69	−0.32 ± 0.40	0.549 ± 0.215
Oura Ring	3.73 ± 1.11	4.46 ± 1.18	5.04 ± 1.90	−2.41 ± 1.07	0.351 ± 0.220
Xmart Ring	6.18 ± 2.23	7.82 ± 3.53	8.93 ± 4.46	+0.26 ± 3.22	0.031 ± 0.108

4.4.1. Huawei Watch Performance

The Huawei Watch demonstrated the highest level of agreement, precision and overall accuracy among all evaluated consumer-grade wearables across the driving protocol:

Bland–Altman Agreement (Figure 9): Exhibited a negligible global systematic bias of −0.3 BPM, with narrow 95% limits of agreement ranging from −8.5 BPM to +7.9 BPM, indicating stable and uniform tracking across various HR intensities and driving dynamics.

Error and Bias Distribution (Figure 10): Recorded the lowest MAE, with a session average of 2.73 ± 1.19 BPM (alongside RMSE of 3.90 ± 1.69 BPM and MAPE of 3.37 ± 1.48%). The session-level systematic bias remained tightly clustered around zero, averaging −0.32 ± 0.40 BPM, confirming minimal systematic offset.

Concordance (Figure 10): Achieved the highest Lin's CCC, with a mean of 0.549 ± 0.215 (reaching up to 0.77 at the 75th percentile), confirming a robust linear association and minimal scale-shift deviation relative to the reference ECG.

4.4.2. Oura Ring Performance

The Oura Ring exhibited moderate tracking accuracy, accompanied by a noticeable underestimation trend during periods of elevated HR:

Bland–Altman Agreement (Figure 9): Displayed a negative global systematic bias of −2.5 BPM, accompanied by wider 95% limits of agreement spanning from −11.6 BPM to +6.6 BPM. The scatter plot reveals increased dispersion and negative divergence during transient physiological elevations.

Error and Bias Distribution (Figure 10): Yielded an average MAE of 3.73 ± 1.11 BPM (with an RMSE of 5.04 ± 1.90 BPM and MAPE of 4.46 ± 1.18%). The systematic bias parameter averaged −2.41 ± 1.07 BPM, confirming a consistent negative shift across the majority of experimental sessions.

Concordance (Figure 10): Recorded a mean Lin's CCC of 0.351 ± 0.220, reflecting moderate concordance constrained by peripheral PPG motion artifacts and vascular perfusion fluctuations inherent to finger-based form factors during active driving maneuvers.

4.4.3. Xmart Ring Performance

The Xmart Ring exhibited the highest error margins, largest performance dispersion, and lowest agreement parameters among the tested devices:

Bland–Altman Agreement (Figure 9): Although the global bias was near zero (+0.3 BPM), the 95% limits of agreement were excessively wide (−18.6 to +19.4 BPM). These extreme deviations are the direct result of a specific software limitation (Xmart ring mobile app, JCRingPro version V1.6.32). The Xmart Ring was intentionally included in the study as a budget-tier commercial device. Originally, the experimental design [63] aimed to utilize its 5-s "workout mode." However, this setting caused critical software instability and continuous Bluetooth drops during initial testing. To maintain the integrity of the driving sessions, the device had to be operated in its default monitoring mode, which restricts data reporting to a coarse 5-min interval. Consequently, the ring holds an outdated HR value steady across prolonged periods, failing entirely to capture the dynamic, transient spikes

that occur during active driving. This 5-min reporting delay is precisely what generates the massive ±20 BPM interpolation errors observed in the analysis.

Error and Bias Distribution (Figure 10): Demonstrated heavily elevated error metrics, recording a mean MAE of 6.18 ± 2.23 BPM, an RMSE of 8.93 ± 4.46 BPM, and an MAPE of 7.82 ± 3.53%. The session bias distribution exhibited high dispersion, averaging +0.26 ± 3.22 BPM with severe outliers reaching near +9 BPM in disturbed segments.

Concordance (Figure 10): Recorded very low concordance, with a mean Lin's CCC approaching near-zero levels (0.031 ± 0.108), demonstrating that raw PPG tracking by this specific device form factor lacks the accuracy required for precise autonomic and cardiac monitoring without advanced real-time motion artifact rejection algorithms.

To further evaluate whether the dynamic tracking fidelity observed in individual case studies extends across the entire cohort, macro-level dynamic HR variability was quantified via the mean standard deviation of HR (SD of HR) across all 19 validated sessions (Figure 11).

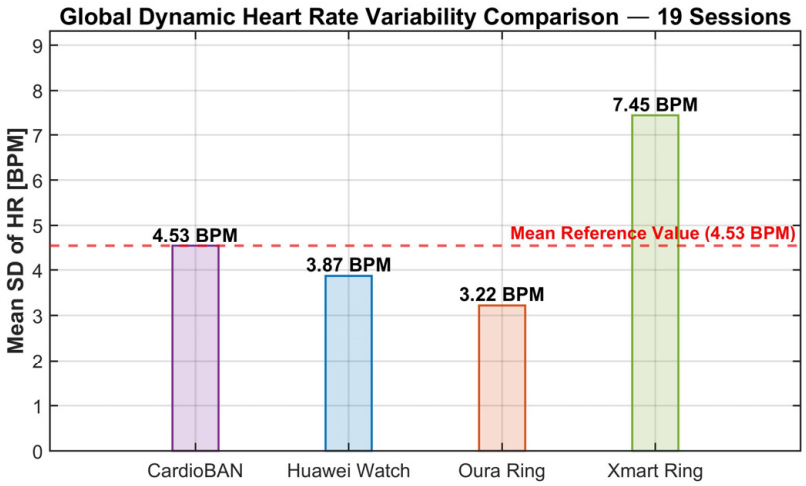


Figure 11. Global comparison of dynamic HR variability, quantified via the mean standard deviation of HR (SD of HR), between the CardioBAN reference standard and consumer wearable devices across the validated cohort (N = 19 sessions).

As illustrated in Figure 11, the CardioBAN reference standard established a mean inherent physiological variability of 4.53 BPM across the cohort. The Huawei Watch closely approximated this dynamic range with a mean SD of HR of 3.87 BPM, confirming its capacity to preserve core autonomic micro-fluctuations during dynamic driving. In contrast, the Oura Ring exhibited an attenuated variability of 3.22 BPM due to aggressive internal smoothing, while the Xmart Ring showed a heavily exaggerated variability of 7.45 BPM, reflecting significant quantization artifacts and false transient spikes induced by uncompensated motion interference. This macro-level variability comparison corroborates that wrist-worn PPG maintains superior accuracy for continuous autonomic tracking, whereas ring-form factors either severely dampen or distort physiological micro-fluctuations under real-world automotive conditions.

To translate the aggregate error distributions into practical reliability metrics for automotive applications, a global device consistency analysis was performed across all 19 validated sessions (Figure 12). This analysis evaluates device success rates against two rigorous criteria: general clinical accuracy and distributional alignment. General clinical accuracy was defined as a session MAE of less than 5 BPM, in accordance with established medical device testing standards (e.g., ANSI/AAMI EC13 guidelines) and recent wearable validation literature [27]. Distributional alignment (the absence of statistically significant

deviation) was defined by a Wilcoxon signed-rank test yielding a $p > 0.05$, indicating no statistically significant distributional deviation from the reference ECG.

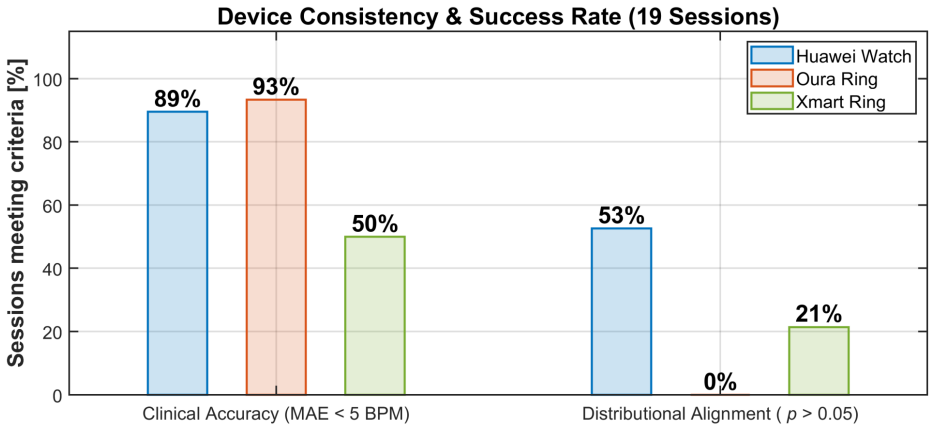


Figure 12. Global device consistency and success rate across the validated cohort ($N = 19$ sessions).

The Huawei Watch demonstrated exceptional reliability, achieving the clinical accuracy threshold in 89% of the driving sessions and maintaining true distributional alignment in 47% of them. This confirms that the smartwatch form factor can reliably acquire high-fidelity, ECG-equivalent data in nearly half of all dynamic driving scenarios.

Notably, while the Oura Ring provided high clinical accuracy in 93% of the sessions, it achieved 0% distributional alignment. This stark discrepancy is a critical finding: it highlights that despite maintaining an acceptable MAE, the ring's proprietary filtering algorithms and hardware limitations introduce continuous systematic deviations. Consequently, the device fails to produce statistically equivalent high-resolution time-series data compared to the reference, rendering it unsuitable for precise autonomic micro-fluctuation analysis without heavy algorithmic compensation. The Xmart Ring exhibited the lowest overall consistency, meeting the clinical accuracy threshold in 50% of the sessions and achieving distributional alignment in 21%.

Ultimately, this consistency analysis demonstrates that evaluating consumer wearables solely on averaged global errors masks underlying tracking deficiencies. While ring-form factors may be sufficient for basic trend monitoring, only wrist-worn PPG has been shown to be reliable enough to consistently match the statistical accuracy of the reference standard during real-world driving.

5. Discussion

The present study sought to definitively evaluate the accuracy and reliability of commercial wearable devices for continuous HR and HRV monitoring in a highly dynamic, real automotive environment. By directly comparing consumer-grade devices against a rigorously denoised, gold-standard ECG, this research assessed their potential to serve as early warning systems for SMI behind the wheel.

5.1. Clinical Suitability for Driver Monitoring and Predictive Viability

The implementation of continuous, real-time physiological monitoring provides a unique opportunity for the early detection of subclinical predictors that precede critical cardiovascular and cerebrovascular conditions. The pathophysiological onset of such events is frequently preceded by an accumulation phase spanning minutes to hours, characterized by extreme sympathetic activation, progressive resting tachycardia, and a severe decline in HRV.

Empirical data from this study confirms the practical existence and measurability of this predictive window within the vehicle cabin. For instance, the reference ECG successfully isolated severe sympathetic crashes (e.g., RMSSD dropping from initial resting levels to under 20 ms within 60-s windows) accompanied by acute tachycardic spikes. Furthermore, in specific driving sessions (e.g., Session 6), the reference pipeline successfully captured dense clusters of ectopic activity (over 130 abrupt inter-beat interval jumps resulting in highly chaotic Poincaré plot dispersion). Crucially, the developed algorithmic pipeline demonstrated a comprehensive capacity to identify bidirectional autonomic extremes. While initial sessions validated the detection of sympathetic tachycardia and stress-induced HRV suppression, subsequent data (e.g., Session 8) confirmed the robust isolation of acute bradycardic transients. Specifically, the reference system successfully quantified profound, instantaneous drops in HR (plummeting to ~63 BPM) coupled with massive parasympathetic surges (RMSSD exceeding 138 ms)—a classic autonomic signature preceding pre-syncopal states and reflex syncope. Furthermore, predictive viability is not strictly limited to extreme tachycardic or bradycardic endpoints. For example, in Session 10, the offline pipeline successfully isolated abrupt autonomic crashes (e.g., an abrupt RMSSD drop to 36.4 ms) even against a backdrop of pronounced resting bradycardia (mean HR < 60 BPM). The capability to detect silent sympathetic arousal during initial resting state is critical for predicting subclinical ischemic events that do not initially manifest with overt macro-tachycardia. Expanding upon these dynamic extremes, data from high-load scenarios characterized by sustained sympathetic storms (e.g., Session 11, exhibiting a resting HR exceeding 90 BPM, an elevated LF/HF ratio of 3.542, and severe parasympathetic collapse with RMSSD values plunging to 13.3 ms) further confirm the robustness of the reference pipeline. The capability to reliably isolate continuous, high-intensity sympathetic overload without signal distortion demonstrates that the pre-syncopal and ischemic accumulation phases remain fully quantifiable even under conditions of severe psychomotor stress and elevated basel metabolic demand. Moreover, extended protocols featuring broad dynamic ranges and reference peaks up to 105 BPM (Session 12) establish absolute temporal integrity across prolonged, high-variability driving intervals. Furthermore, continuous severe tachycardic regimes characterized by persistent HR elevations exceeding 103 BPM coupled with sustained autonomic compression (SDNN plateauing below 20 ms, Session 14) confirm the pipeline's capability to capture prolonged psychophysiological overload without saturation. In addition, complex spectral profiles featuring extreme sympathetic dominance coupled with transient dynamic surges (Session 16, exhibiting an elevated LF/HF ratio of 2.692 alongside sharp tachycardic peaks reaching 115 BPM) further validate the pipeline's advanced capacity to isolate multi-domain autonomic distress markers under real-world driving stressors. This validates that automotive environments do not preclude the detection of critical pre-syncopal, ischemic, or arrhythmic markers, provided the sensor possesses the necessary high-precision temporal resolution.

Crucially, this comprehensive validation is not derived from a single isolated scenario but is systematically corroborated across the entire spectrum of the experimental cohort. In-depth, session-by-session analyses—ranging from extreme tachycardic sympathetic storms (e.g., Sessions 17 and 20) to prolonged states of stable parasympathetic dominance and subsequent autonomic transitions (e.g., Sessions 18, 19, and 21)—consistently reinforce the primary findings. Regardless of whether the driver exhibited a severe, protracted autonomic crisis or a normal physiological state, the reference pipeline invariably maintained its structural integrity, while the tested wearable form factors exhibited the exact same stratification in performance constraints. This session-level consistency solidifies the conclusion that the observed technological and physiological limitations are universal phenomena within the automotive environment, rather than statistical anomalies.

5.2. Form-Factor Influence on Accuracy

The dramatic difference in performance observed between the wrist-worn smartwatch and the finger-worn smart rings can be directly attributed to the physical and biomechanical realities of the vehicle cabin. As outlined in Section 1.2, operating the steering wheel requires continuous isometric tension in the forearms and fingers, which dynamically alters local hydrostatic pressure. This phenomenon (presented in Figure 13) was empirically demonstrated by the severe quantization artifacts (banding) and prolonged flatlining observed in the Bland–Altman and time-series analyses of the tested ring form-factors, indicating a complete loss of dynamic range during steering. This interference was confirmed by cross-referencing the degraded signal timestamps with our synchronized video logs, which clearly showed that ring signal flatlining consistently coincided with prolonged, tense gripping of the steering wheel.

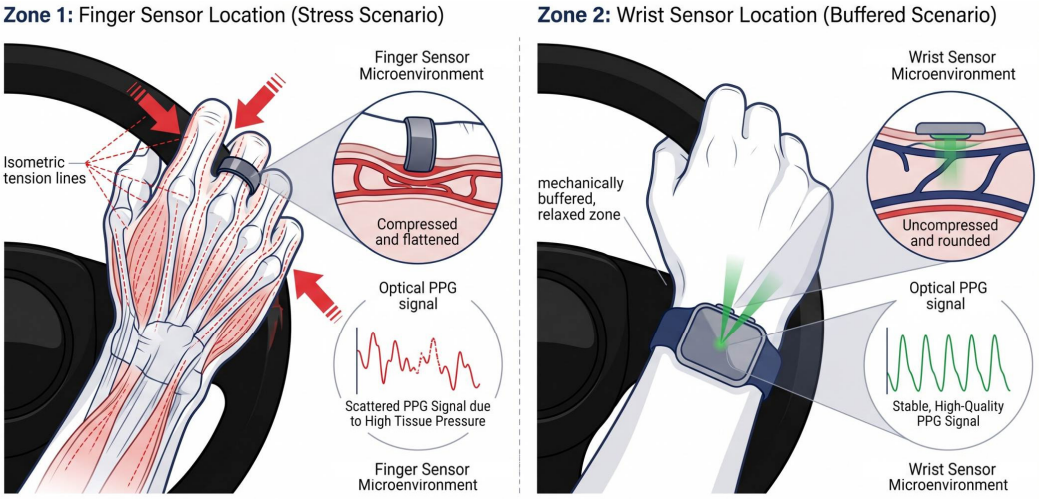


Figure 13. Conceptual schematic illustrating the biomechanical limitations of ring-form factors during driving. (**left side**) Zone 1 (Finger) is subjected to continuous isometric tension and direct grip pressure against the steering wheel, causing capillary compression and severe PPG perfusion decoupling. In contrast, (**right side**) Zone 2 (Wrist) remains mechanically buffered, ensuring stable optical contact and reliable tracking via smartwatches. Red arrows indicate the direction of direct mechanical pressure against the steering wheel. Red dashed lines represent the zones of continuous isometric muscular tension. In the magnified microenvironment panels, red and blue tubes represent arteries and veins respectively, while the green shaded beam in Zone2 represents the optical PPG sensor's light emission. Solid green and scattered (dashed) red line graphs depict high-quality and degraded PPG signal waveforms, respectively.

Crucially, the biomechanical limitations of the ring form-factor not only result in signal loss but can generate dangerous false-positive artifacts. For instance, in Session 10, the isometric pressure on the steering wheel during steady bradycardic driving forced the Xmart ring to start coarse interpolation, resulting in massive artificial spikes (exceeding 110 BPM) while the actual physiological rate remained near 60 BPM. In an autonomous intervention scenario, such mechanically induced hallucinations could trigger unwarranted emergency vehicle braking, endangering traffic flow.

Furthermore, high-frequency vibrations from the road surface and the engine transmit directly through the steering column to the driver's gripping fingers, heavily contaminating the optical signal. The wrist, while still subject to the general kinetic movement of steering, is relatively buffered from the direct, intense isometric squeezing applied by the fingers gripping the wheel. Consequently, the smartwatch maintained superior optical contact and signal stability, which explains its significantly lower MAE and higher distributional alignment compared to the ring form factors. Even during intense autonomic shifts and

tachycardic spikes, the wrist-worn form factor maintained consistent error distributions and high concordance (Lin's CCC > 0.65). This stability was similarly robust across low resting HR, with wrist-PPG achieving distributional alignment ($p > 0.67$) to the reference signal during prolonged parasympathetic dominance. Similarly, under conditions of maximal physiological stress and elevated driver kinetics (Session 11), the wrist-worn smartwatch maintained distributional alignment with the reference standard ($p = 0.0853$), outperforming the ring form-factors that suffered from extensive negative proportional bias and clipping. This linear tracking precision was robustly sustained across extended dynamic modes (Session 12, $p = 0.4326$), further highlighting the vulnerability of ring sensors to perfusion decoupling. This exceptional accuracy was further reinforced during prolonged, high-intensity continuous tachycardia (Session 14, Mean HR = 103.3 BPM), where the wrist-worn smartwatch achieved superior precision with a minimal MAE of 1.18 BPM and tight agreement limits. This superior tracking performance of the wrist-form factor was robustly maintained during acute high-intensity transients (Session 16, achieving a high Lin's CCC of 0.776 and a minimal global bias of −0.35 BPM), confirming its structural suitability for continuous in-vehicle monitoring. Fundamentally, the combination of constant isometric pressure and direct vibration transmission renders the anatomical location of the phalanx biomechanically incompatible with reliable PPG monitoring during driving, regardless of the device's hardware tier.

The results of this study reveal a critical distinction in the ability of consumer wearables to capture these slowly accumulating trends. For tracking macroscopic physiological shifts, such as a gradual escalation in mean HR, commercial devices exhibit substantial proficiency. Both the wrist-worn smartwatch and the finger-worn smart rings achieved high clinical accuracy regarding mean HR, indicating their potential reliability for identifying broad, slow-onset physiological deteriorations. However, the most definitive precursors to ischemic events or arrhythmias reside within the fine, chaotic millisecond variations in the inter-beat intervals, rather than the average pulse.

The data unequivocally indicates that relying on certain form factors to capture these nuanced autonomic degradations remains highly dangerous. Specifically, the tested smart rings failed entirely to produce statistically equivalent high-resolution time-series data compared to the reference, achieving 0% distributional alignment. Their aggressive internal smoothing erases the vital micro-fluctuations necessary for tracking the gradual decline in HRV over time. Conversely, the wrist-worn smartwatch successfully preserved these core autonomic micro-fluctuations, maintaining true distributional alignment in 47% of the driving sessions. This demonstrates that wrist-based PPG possesses a significantly higher real-world potential not only for tracking basic HR trends but also for capturing the vital autonomic degradation required for the early prediction of sudden incapacitation during the critical symptom accumulation phase.

5.3. The Black-Box Algorithm Problem

A critical finding of this study is the highly misleading nature of global error metrics when applied to commercial devices that utilize proprietary, "black-box" software algorithms. Most commercial devices attempt to overcome poor PPG signal quality during motion by aggressive smoothing, moving averages, or data imputation for missing segments. This aggressive internal filtering artificially distorts the raw physiological data, prioritizing a visually stable average HR over true beat-to-beat accuracy.

For instance, the Oura Ring provided high clinical accuracy (MAE < 5 BPM) in 93% of the sessions, yet it achieved 0% distributional alignment. Bland–Altman analysis revealed a pronounced negative proportional bias, indicating that the device systematically underestimates the HR specifically during acute physiological spikes (e.g., clipping real values of

>90 BPM down to 80–85 BPM). This asymmetric degradation implies that such devices fail precisely when their accuracy is most critical—during the onset of physiological stress.

This stark discrepancy highlights that while the average pulse appears accurate, the device's algorithms completely obliterate the microsecond variability (HRV). Moreover, these outlier-removal algorithms are inherently designed to suppress sudden data jumps, meaning they actively erase genuine ectopic beats (PVCs/PACs), misclassifying critical predictors of imminent arrhythmias as mere motion artifacts (Figure 14).

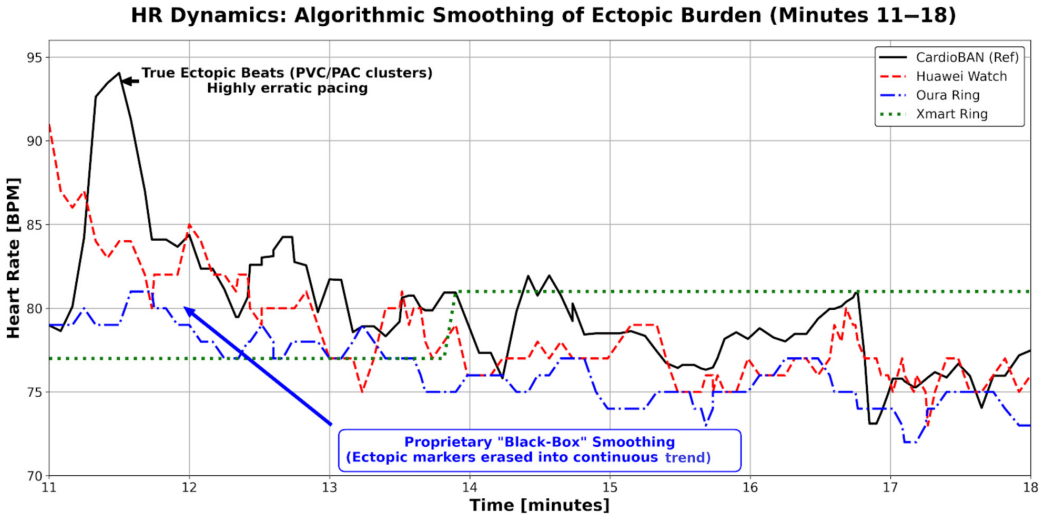


Figure 14. Visualization of "black-box" algorithmic smoothing obliterating critical physiological micro-variability during high ectopic burden (Session 6). The solid black line (Reference ECG) captures highly erratic pacing and clusters of true ectopic activity. In stark contrast, the proprietary algorithms of finger-worn devices (e.g., the dash–dot blue line of the Oura Ring and the dotted green line of the Xmart Ring) misclassify these ectopic markers as motion artifact, artificially erasing them into a continuous, falsely stable trend.

This was explicitly proven during episodes of high ectopic burden, where the medical reference identified substantial erratic pacing, yet the commercial smart ring aggressively smoothed the signal into a continuous, falsely stable trend—the blue line in Figure 14. Similarly, during acute bradycardic events, aggressive signal smoothing creates dangerous latencies, artificially maintaining higher pulse readings while the true physiological state deteriorates rapidly. Furthermore, "black-box" filtering compresses dynamic range indiscriminately across the physiological spectrum. Even during resting bradycardia, in the absence of gross motor interference (Session 10), proprietary smoothing visibly flattened the micro-fluctuations necessary to calculate RMSSD (Figure 15), demonstrating that the absence of a transparent SQA allows artificial signal compression to masquerade as valid physiological stability.

Moreover, during acute tachycardic surges exceeding 110 BPM (Session 11), proprietary ring algorithms manifested severe negative proportional scaling, systematically clipping peak responses and failing to resolve the compressed short-term variability (SD_1 = 19.5 ms). This saturation trend was corroborated during high-load peaks reaching 105 BPM (Session 12), where ring algorithms systematically attenuated transient surges. This limitation was further underscored during extreme autonomic surges (Session 16, where proprietary finger-ring algorithms suffered from severe amplitude clipping, flattening an acute 115 BPM tachycardic peak down to ~95 BPM as illustrated in Figure 16).

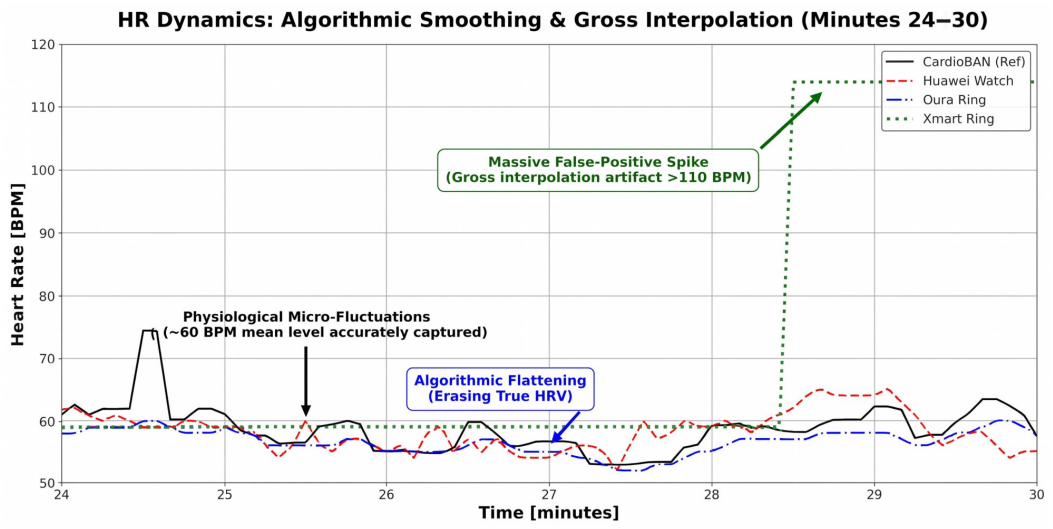


Figure 15. Impact of proprietary algorithmic smoothing and gross interpolation during resting bradycardia (Session 10). The reference ECG (solid black line) captures natural physiological micro-fluctuations around 60 BPM. In contrast, ring form-factors not only artificially flatten this essential variability (thereby compromising RMSSD calculations) but also introduce massive, false-positive tachycardic spikes (exceeding 110 BPM, dotted green line) due to motion-induced signal loss and aggressive interpolation.

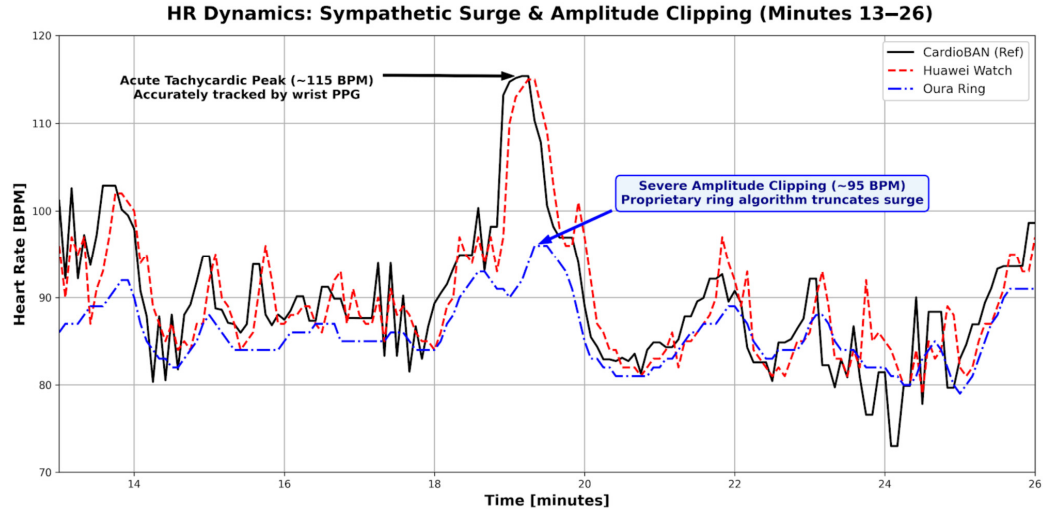


Figure 16. Severe amplitude clipping during an acute sympathetic surge (Session 16). While the reference ECG (solid black line) accurately tracks a rapid tachycardic transient peaking at approximately 115 BPM, the proprietary algorithms of finger-worn device (blue dash–dot line) systematically attenuate this physiological spike. This saturation effect arbitrarily flattens the autonomic response down to ~95 BPM, effectively masking a critical marker of physiological distress.

This systematic attenuation confirms that closed-source filtering structures inherently disable consumer devices from registering critical autonomic distress. Healthy physiological function is characterized by this subtle, chaotic variability. Consequently, the sudden absence of these micro-fluctuations acts as a vital clinical predictor for an impending infarction, stroke, or severe arrhythmia. By masking these critical physiological micro-fluctuations, closed-source smoothing algorithms create a dangerous false sense of security, rendering the heavily processed PRV signals inadequate for reliable integration into autonomous vehicle safety interventions.

5.4. Study Limitations and Future Directions

Despite the rigorous experimental design and the high-fidelity offline pipeline established in this research, certain limitations must be acknowledged. The current validation framework was executed with a cohort of 21 healthy participants. While this sample size provides statistically robust evidence regarding the technological and biomechanical constraints of consumer wearables during driving, it does not fully encompass the physiological heterogeneity of the general driving population. Future research must scale this methodology across larger, more diverse cohorts, specifically incorporating individuals with previously diagnosed cardiovascular pathologies, such as chronic arrhythmias or ischemic heart disease. Evaluating these systems within clinical cohorts under real-world driving conditions will be a critical next step to definitively validate the predictive capability of unobtrusive monitoring systems for drivers at the highest risk of SMI.

Furthermore, the significant data loss we observed in these disrupted sessions (Sessions 5 and 7) points to a practical hurdle for naturalistic driving research. Even when reference sensors were securely attached just below the chest, they remained vulnerable to everyday human behaviors. Actions like expressive gesturing, shifting posture due to fatigue, and erratic steering make standard attachment methods prone to mechanical decoupling. Since the applied automated algorithms successfully catch and reject these corrupted segments, the real challenge for future studies is preventing the physical signal degradation at its source. To ensure uninterrupted data collection, researchers need to optimize the physical sensor setup to be universally adaptable for any user. This might involve exploring alternative placements to avoid seatbelt friction, using stronger medical-grade adhesives, or employing compressive garments to stabilize soft tissue during dynamic driving. Crucially, the chosen hardware and attachment methods must encompass the full spectrum of human diversity. The physical setup must accommodate everyone from very lean to heavily built individuals, while also accounting for drivers with highly temperamental or expressive habits. Meeting these diverse needs is essential to guarantee stable contact and reliable readings regardless of body type or driving style.

6. Conclusions

This study implemented a high-precision offline validation framework to evaluate the measurement accuracy and reliability of commercial wearable devices for continuous physiological driver monitoring in dynamic vehicle environments. By benchmarking wrist- and finger-based PPG against an artifact-suppressed clinical electrocardiographic reference, the investigation identified fundamental technological and biomechanical constraints.

The empirical data indicate that wrist-worn smartwatches maintain high measurement stability, meeting clinical accuracy criteria in 89% of driving sessions and distributional alignment in 47%. These characteristics render wrist-form modalities viable for tracking subclinical autonomic variations during driving. Conversely, ring-form factors are structurally constrained by the sustained isometric tension and high-frequency vibration noise characteristic of steering maneuvers, resulting in acute perfusion decoupling and 0% distributional alignment. Furthermore, the reliance of consumer devices on proprietary closed-source filtering architectures systematically attenuates the high-frequency micro-fluctuations (HRV) essential for detecting acute autonomic decompensation and arrhythmogenesis.

Ultimately, while commercial smartwatches provide a scalable hardware foundation for in-cabin telemetry, advanced predictive safety frameworks must circumvent intrinsic smoothing heuristics. Effective integration into life-saving automotive interventions requires access to raw high-resolution data coupled with transparent, application-specific signal processing algorithms.

Author Contributions: Conceptualization, G.P. and H.R.; methodology, H.R.; formal analysis H.R.; investigation, H.R.; resources, G.P. and H.R.; writing—original draft preparation, H.R.; writing—review and editing, G.P.; visualization, H.R.; project administration, G.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been supported by the European Regional Development Fund within the OP "Research, Innovation and Digitalization Programme for Intelligent Transformation 2021–2027". Project No. BG16RFPR002-1.014-0005 Center of competence "Smart Mechatronics, Eco-and Energy Saving Systems and Technologies".

Institutional Review Board Statement: This study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee for Research Involving Human Participants of the Technical University of Sofia (Protocol No. 4/20 October 2025).

Informed Consent Statement: Informed consent was obtained from all subjects involved in this study.

Data Availability Statement: The raw data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

HR	Heart Rate
HRV	Heart Rate Variability
SDNN	Standard Deviation of Normal-to-Normal intervals
RMSSD	Root Mean Square of Successive Differences
AFib	Atrial Fibrillation
PVC	Premature Ventricular Contractions
PAC	Premature Atrial Contractions
RR	R–R interval (Interval between consecutive R–peaks)
LF	Low Frequency
PR	Pulse Rate
SMI	Sudden Medical Illnesses
HF	High Frequency
PPG	Photopletysmography
PTT	Pulse Transit Time
PRV	Pulse Rate Variability
BW	Baseline Wander
EMG	Electromyography
SQA	Signal Quality Assessment
MAD	Median Absolute Deviation
CCC	Concordance Correlation Coefficient
BPM	Beats Per Minute
IMU	Inertial Measurement Unit
IIR	Infinite Impulse Response
FIR	Finite Impulse Response
QRS	Q-R-S complex of ECG signal
MODWT	Maximal Overlap Discrete Wavelet Transform
ALDTF	Adaptive Layer-Dependent Thresholding Function
GPS	Global Positioning System
DC	Direct Current
SD	Standard Deviation
NZOPP	Non-Zero Periodic Peak
IHR	Instantaneous Heart Rate
NaN	Not a Number

MAE	Mean Absolute Error
RMSE	Root Mean Square Error
MAPE	Mean Absolute Percentage Error
BMI	Body Mass Index
IBI	Inter-beat Interval

References

1. Rahman, A.; Barber, F.; Lovric, M.; Woodman, R.; Donzella, V. Driver Monitoring Systems: Advances, Challenges, and Future Directions in Noncontact Sensor Technologies: A Review. *IEEE Sens. J.* **2026**, *26*, 3540–3557. [CrossRef]
2. Petrova, G.; Radev, H.; Shopov, M.; Kakanakov, N. A Taxonomy of Methods, Techniques and Sensors for Acquisition of Physiological Signals in Driver Monitoring Systems. *Appl. Sci.* **2025**, *15*, 7609. [CrossRef]
3. Phipps, J.; Passage, B.; Sel, K.; Martinez, J.; Saadat, M.; Koker, T.; Damaso, N.; Davis, S.; Palmer, J.; Claypool, K.; et al. Early adverse physiological event detection using commercial wearables: Challenges and opportunities. *npj Digit. Med.* **2024**, *7*, 136. [CrossRef] [PubMed]
4. Miao, Q.; Zhang, Y.-L.; Miao, Q.-F.; Yang, X.-A.; Zhang, F.; Yu, Y.-G.; Li, D.-R. Sudden Death from Ischemic Heart Disease While Driving: Cardiac Pathology, Clinical Characteristics, and Countermeasures. *Med. Sci. Monit.* **2021**, *27*, e929212. [CrossRef] [PubMed]
5. Liu, S.; Cui, Y.; Chen, M. Heart rate variability: A multidimensional perspective from physiological marker to brain-heart axis disorders prediction. *Front. Cardiovasc. Med.* **2025**, *12*, 1630668. [CrossRef] [PubMed]
6. Fatih Gündüz, A.; Fatih Talu, M. Atrial fibrillation classification and detection from ECG recordings. *Biomed. Signal Process. Control* **2023**, *82*, 104531. [CrossRef]
7. Moelgg, K.; Karisik, A.; Buergi, L.; Scherer, L.; Delazer, L.; Dejakum, B.; Felicetti, S.; Koehler, T.; Granna, J.; Boehme, C.; et al. Value of routine heart rate variability parameters for atrial fibrillation detection in ischaemic stroke and high-risk TIA patients. *Front. Stroke* **2026**, *4*, 1727719. [CrossRef] [PubMed]
8. Meinel, T.; Arnold, M.; Roten, L.; Krisai, P.; Mono, M.-L.; Gebhart, C.; Bonati, L.; Kahles, T.; Fischer, U.; Arnold, M.; et al. Swiss Stroke Society position paper on atrial fibrillation monitoring and management after ischaemic stroke: A shift from understanding the index stroke to preventing the next one. *Swiss Med. Wkly.* **2025**, *155*, 4170. [CrossRef] [PubMed]
9. Gao, S.; Tang, X. Autonomic Nervous Dysfunction and Ultra-Short-Term Heart Rate Variability in Atrial Fibrillation: Recent Advances in Early Detection. *J. Cardiovasc. Dev. Dis.* **2026**, *13*, 286. [CrossRef] [PubMed]
10. Sutton, R.; Kenny, R.A.; Benditt, D.G. Syncope: Advances in Diagnosis and Treatment 2024. *J. Cardiovasc. Electrophysiol.* **2025**, *36*, 2671–2676. [CrossRef] [PubMed]
11. Hege, A.; Lemke, M.K.; Apostolopoulos, Y.; Sönmez, S. Occupational health disparities among U.S. long-haul truck drivers: The influence of work organization and sleep on cardiovascular and metabolic disease risk. *PLoS ONE* **2018**, *13*, e0207322. [CrossRef] [PubMed]
12. Choi, J.; Cha, W.; Park, M.-G. Declining Trends of Heart Rate Variability According to Aging in Healthy Asian Adults. *Front. Aging Neurosci.* **2020**, *12*, 610626. [CrossRef] [PubMed]
13. Kaze, A.D.; Fonarow, G.C.; Echouffo-Tcheugui, J.B. Cardiac Autonomic Dysfunction and Risk of Silent Myocardial Infarction Among Adults with Type 2 Diabetes. *J. Am. Heart Assoc.* **2023**, *12*, e029814. [CrossRef] [PubMed]
14. Gospodinova, E. Methods of Nonlinear Dynamics for Heart Rate Variability Analysis. *Innov. STEM Educ.* **2022**, *4*, 32–38. [CrossRef]
15. Liang, T.; Yilmaz, G.; Soon, C.-S. Deriving Accurate Nocturnal Heart Rate, rMSSD and Frequency HRV from the Oura Ring. *Sensors* **2024**, *24*, 7475. [CrossRef] [PubMed]
16. Nazir, A.; Nazir, A.; Shah Wali Jamal, M.; Sadiq, S.U.R.; Aman, S.; Mustapha, M.J.; Lawal, S.O.; AbdulKareem, M.O.; Bamigbola, M.F. Wearable Technology and Its Potential Role in Cardiovascular Health Monitoring and Disease Management. *Health Sci. Rep.* **2025**, *8*, e71486. [CrossRef] [PubMed]
17. Wessel, N.; Schumann, A.; Schirdewan, A.; Voss, A.; Kurths, J. Entropy measures in heart rate variability data. In *Medical Data Analysis*; Springer: Berlin/Heidelberg, Germany, 2000; pp. 78–87. [CrossRef] [PubMed]
18. Baiense, J.P.; Coelho, P.J.; Pires, I.M.; Velez, F.J. Wearable solution for health monitoring of car drivers. *Procedia Comput. Sci.* **2024**, *238*, 73–80. [CrossRef]
19. Soumma, S.B.; Mamun, A.; Ghasemzadeh, H. AI-Powered Wearable Sensors for Health Monitoring and Clinical Decision Making. *Curr. Opin. Biomed. Eng.* **2025**, *36*, 100628. [CrossRef]

20. Mejía-Mejía, E.; Budidha, K.; Abay, T.Y.; May, J.M.; Kyriacou, P.A. Heart Rate Variability (HRV) and Pulse Rate Variability (PRV) for the Assessment of Autonomic Responses. *Front. Physiol.* **2020**, *11*, 779. [CrossRef] [PubMed]
21. Elgendi, M.; Fletcher, R.; Liang, Y.; Howard, N.; Lovell, N.H.; Abbott, D.; Lim, K.; Ward, R. The use of photoplethysmography for assessing hypertension. *npj Digit. Med.* **2019**, *2*, 60. [CrossRef] [PubMed]
22. Schäfer, A.; Vagedes, J. How accurate is pulse rate variability as an estimate of heart rate variability? A review on studies comparing photoplethysmographic technology with an electrocardiogram. *Int. J. Cardiol.* **2013**, *166*, 15–29. [CrossRef] [PubMed]
23. Gil, E.; Orini, M.; Bailón, R.; Vergara, J.M.; Mainardi, L.; Laguna, P. Photoplethysmography pulse rate variability as a surrogate measurement of heart rate variability during non-stationary conditions. *Physiol. Meas.* **2010**, *31*, 1271–1290. [CrossRef] [PubMed]
24. Hernandez, R.; Schneider, S.; De Vries, H.J.; Fanning, J.; Ehrmann, D.; Jin, H.; Moore, R.C.; Juengst, S.; Striegel, A.; Ginsberg, J.P.; et al. Resting Heart Rate Variability Measured by Consumer Wearables and Its Associations with Diverse Health Domains in Five Longitudinal Studies. *Sensors* **2025**, *25*, 7147. [CrossRef] [PubMed]
25. Zhao, Y.; Chen, P.; Zhang, Y.; Huo, S.; Yu, D.; Zeng, X.; Zhang, W. Heart rate variability and its modulation by nutrients: A narrative review on implications for cardiovascular aging. *Front. Neurosci.* **2025**, *19*, 1654796. [CrossRef] [PubMed]
26. Doherty, C.; Baldwin, M.; Keogh, A.; Caulfield, B.; Argent, R. Keeping Pace with Wearables: A Living Umbrella Review of Systematic Reviews Evaluating the Accuracy of Consumer Wearable Technologies in Health Measurement. *Sports Med.* **2024**, *54*, 2907–2926. [CrossRef] [PubMed]
27. Gruden, T.; Stojmenova, K.; Sodnik, J.; Jakus, G. Assessing Drivers' Physiological Responses Using Consumer Grade Devices. *Appl. Sci.* **2019**, *9*, 5353. [CrossRef]
28. Kantrowitz, A.B.; Ben-David, K.; Morris, M.; Wittels, H.L.; Wishon, M.J.; McDonald, S.M.; Renaghan, E.J.; Feigenbaum, L.A.; Wittels, S.H. Pulse rate variability is not the same as heart rate variability: Findings from a large, diverse clinical population study. *Front. Physiol.* **2025**, *16*, 1630032. [CrossRef] [PubMed]
29. Georgiou, K.; Larentzakis, A.V.; Khamis, N.N.; Alsuhaibani, G.I.; Alaska, Y.A.; Giallafos, E.J. Can Wearable Devices Accurately Measure Heart Rate Variability? A Systematic Review. *Folia Med.* **2018**, *60*, 7–20. [CrossRef] [PubMed]
30. Ranjan, S.; Baria, D.; Reddy, Y.C. Systematic Review Wearable Technology for Cardiac Rhythm Monitoring. *J. Heart Valve Dis.* **2025**, *30*, 71–77.
31. Burma, J.S.; Griffiths, J.K.; Lapointe, A.P.; Oni, I.K.; Soroush, A.; Carere, J.; Smirl, J.D.; Dunn, J.F. Heart Rate Variability and Pulse Rate Variability: Do Anatomical Location and Sampling Rate Matter? *Sensors* **2024**, *24*, 2048. [CrossRef] [PubMed]
32. Allen, J. Photoplethysmography and its application in clinical physiological measurement. *Physiol. Meas.* **2007**, *28*, R1–R39. [CrossRef] [PubMed]
33. Chan, G.; Cooper, R.; Hosanee, M.; Welykholowa, K.; Kyriacou, P.A.; Zheng, D.; Allen, J.; Abbott, D.; Lovell, N.H.; Fletcher, R.; et al. Multi-Site Photoplethysmography Technology for Blood Pressure Assessment: Challenges and Recommendations. *J. Clin. Med.* **2019**, *8*, 1827. [CrossRef] [PubMed]
34. Mukkamala, R.; Hahn, J.-O.; Inan, O.T.; Mestha, L.K.; Kim, C.-S.; Toreyin, H.; Kyal, S. Toward Ubiquitous Blood Pressure Monitoring via Pulse Transit Time: Theory and Practice. *IEEE Trans. Biomed. Eng.* **2015**, *62*, 1879–1901. [CrossRef] [PubMed]
35. Ziyadidegan, S.; Sadeghi, N.; Razavi, M.; Baharlouei, E.; Janfaza, V.; Kazeminasab, S.; Pesarakli, H.; Javid, A.H.; Sasangohar, F. Quantifying Mental Stress Using Cardiovascular Responses: A Scoping Review. *Sensors* **2025**, *25*, 4281. [CrossRef] [PubMed]
36. Liu, B.; Zhang, Z.; Di, X.; Wang, X.; Xie, L.; Xie, W.; Zhang, J. The Assessment of Autonomic Nervous System Activity Based on Photoplethysmography in Healthy Young Men. *Front. Physiol.* **2021**, *12*, 733264. [CrossRef] [PubMed]
37. Radev, H.; Petrova, G.; Spasov, G. Driver physiological parameters monitoring–initial study in real-road driving. In *2024 XXXIII International Scientific Conference Electronics (ET)*; IEEE: Piscataway, NJ, USA, 2024; pp. 1–6. [CrossRef]
38. Babusiak, B.; Smondrk, M.; Trpis, L.; Gajdosik, T.; Madaj, R.; Gajdac, I. Smart Steering Wheel Prototype for In-Vehicle Vital Sign Monitoring. *Sensors* **2026**, *26*, 477. [CrossRef] [PubMed]
39. AlArnaout, Z.; Zaki, C.; Kotb, Y.; AlAkkoumi, M.; Mostafa, N. Exploiting heart rate variability for driver drowsiness detection using wearable sensors and machine learning. *Sci. Rep.* **2025**, *15*, 24898. [CrossRef] [PubMed]
40. Montanari, A.; Ferlini, A.; Balaji, A.N.; Mascolo, C.; Kawsar, F. EarSet: A Multi-Modal Dataset for Studying the Impact of Head and Facial Movements on In-Ear PPG Signals. *Sci. Data* **2023**, *10*, 850. [CrossRef] [PubMed]
41. Boboc, R.G.; Butilă, E.V.; Butnariu, S. Leveraging Wearable Sensors in Virtual Reality Driving Simulators: A Review of Techniques and Applications. *Sensors* **2024**, *24*, 4417. [CrossRef] [PubMed]
42. Germini, F.; Noronha, N.; Borg Debono, V.; Abraham Philip, B.; Pete, D.; Navarro, T.; Keepanasseril, A.; Parpia, S.; De Wit, K.; Iorio, A. Accuracy and Acceptability of Wrist-Wearable Activity-Tracking Devices: Systematic Review of the Literature. *J. Med. Internet Res.* **2022**, *24*, e30791. [CrossRef] [PubMed]

43. Stergiou, G.S.; Mukkamala, R.; Avolio, A.; Kyriakoulis, K.G.; Mieke, S.; Murray, A.; Parati, G.; Schutte, A.E.; Sharman, J.E.; Asmar, R.; et al. Cuffless blood pressure measuring devices: Review and statement by the European Society of Hypertension Working Group on Blood Pressure Monitoring and Cardiovascular Variability. *J. Hypertens.* **2022**, *40*, 1449–1460. [CrossRef] [PubMed]
44. Crooks, L.; van Schoor, G.M.; Everhart, A.; Byrom, B.; Sondag, T. Validity, Reliability, and Regulatory Considerations for Consumer Devices in Clinical Research Data Collection. *Appl. Clin. Trials* **2026**. Available online: https://www.appliedclinicaltrialsonline.com/view/validity-reliability-regulatory-considerations-consumer-devices-clinical-research-data-collection (accessed on 29 August 2026).
45. Dial, M.B.; Hollander, M.E.; Vatne, E.A.; Emerson, A.M.; Edwards, N.A.; Hagen, J.A. Validation of nocturnal resting heart rate and heart rate variability in consumer wearables. *Physiol. Rep.* **2025**, *13*, e70527. [CrossRef] [PubMed]
46. Hosanee, M.; Chan, G.; Welykholowa, K.; Cooper, R.; Kyriacou, P.A.; Zheng, D.; Allen, J.; Abbott, D.; Menon, C.; Lovell, N.H.; et al. Cuffless Single-Site Photoplethysmography for Blood Pressure Monitoring. *J. Clin. Med.* **2020**, *9*, 723. [CrossRef] [PubMed]
47. Carrier, B.; Bunn, J.A.; Eschbach, C.; Reece, J.D.; Welk, G.J.; Dolezal, B.A.; Navalta, J.W. The WEAR-BOT checklist: A risk of bias tool for evaluating validity and reliability research in wearable technology. *PLoS ONE* **2026**, *21*, e0338014. Correction in *PLoS ONE* **2026**, *21*, e0345217. https://doi.org/10.1371/journal.pone.0345217. [CrossRef] [PubMed]
48. Corso, J.L.; Peikon, E. The Challenge of Measuring Exercise: Advancing Metrological Barriers in Wearable Sensing. *JMIR mHealth uHealth* **2025**, *13*, e79347. [CrossRef] [PubMed]
49. Paluch, K.; Szypuła, Z.; Nowak, A.; Bochenek, O.; Koper, M.; Kałuża, J.; Konaszczuk, A.; Ratyna, K.; Kozyra, O.; Skarbek, M. Accuracy of Wrist-Worn Heart Rate Monitors: A Comprehensive Review of Smartwatches in Exercise Monitoring. *Qual. Sport* **2024**, *32*, 55350. [CrossRef]
50. Pecik, M.; Vavrinsky, E.; Vitazkova, D.; Kosnacova, H.; Nevrela, J.; Foltan, E. Respiratory Monitoring in Motion: An Overview of Wearable Methods and Algorithmic Approaches for Reliable Assessment. *Biosensors* **2026**, *16*, 306. [CrossRef] [PubMed]
51. Hu, X.; Sgherza, T.R.; Nothrup, J.B.; Fresco, D.M.; Naragon-Gainey, K.; Bylsma, L.M. From lab to life: Evaluating the reliability and validity of psychophysiological data from wearable devices in laboratory and ambulatory settings. *Behav. Res. Methods* **2024**, *56*, 1–20. [CrossRef] [PubMed]
52. Choe, J.-P.; Oh, S.-E.; Kang, M. Accuracy of Polar Wrist-Worn Devices Compared to Chest-Worn Devices: A Bland-Altman Meta-Analysis. *Meas. Phys. Educ. Exerc. Sci.* **2026**, 1–11. [CrossRef]
53. Hahnen, C.; Freeman, C.G.; Haldar, N.; Hamati, J.N.; Bard, D.M.; Murali, V.; Merli, G.J.; Joseph, J.I.; Van Helmond, N. Accuracy of Vital Signs Measurements by a Smartwatch and a Portable Health Device: Validation Study. *JMIR mHealth uHealth* **2020**, *8*, e16811. [CrossRef] [PubMed]
54. Vermunicht, P.; Makayed, K.; Buyck, C.; Knaepen, L.; Piedrahita Giraldo, J.S.; Naessens, S.; Hens, W.; Craenenbroeck, E.V.; Laukens, K.; Desteghe, L.; et al. Continuous heart rate measurements in patients with cardiac disease: Device comparison and development of a novel artefact removal procedure. *Digit. Health* **2025**, *11*, 20552076251337598. [CrossRef] [PubMed]
55. Altamira-Colado, E.; Bravo-Zanoguera, M.; Cuevas-González, D.; Reyna-Carranza, M.; López-Avitia, R. Removal of ECG baseline wander using a resting cycle template. In Proceedings of the 8th International Electronic Conference on Sensors and Applications, Virtual, 1–15 November 2021. [CrossRef]
56. Warnecke, J.M.; Lasenby, J.; Deserno, T.M. Robust in-vehicle heartbeat detection using multimodal signal fusion. *Sci. Rep.* **2023**, *13*, 20864. Correction in *Sci. Rep.* **2024**, *14*, 6224. https://doi.org/10.1038/s41598-024-56883-9. [CrossRef] [PubMed]
57. Ding, J.; Tang, Y.; Chang, R.; Li, Y.; Zhang, L.; Yan, F. Reduction in the Motion Artifacts in Noncontact ECG Measurements Using a Novel Designed Electrode Structure. *Sensors* **2023**, *23*, 956. [CrossRef] [PubMed]
58. Adams, B. High-Reliability Signal Quality Validation for Biosignals Using Sensor Fusion and Software Indices. *Sensors* **2026**, *26*, 3478. [CrossRef] [PubMed]
59. Dobrev, D.; Neycheva, T.; Krasteva, V.; Jekova, I. Design of High-Pass and Low-Pass Active Inverse Filters to Compensate for Distortions in RC-Filtered Electrocardiograms. *Technologies* **2025**, *13*, 159. [CrossRef]
60. Zhang, Y.; Yu, K.; Huang, C.; Qu, R.; Fan, Z.; Zhu, P.; Chen, W.; Hao, J. Adaptive Layer-Dependent Threshold Function for Wavelet Denoising of ECG and Multimode Fiber Cardiorespiratory Signals. *Sensors* **2025**, *25*, 7644. [CrossRef] [PubMed]
61. Xiao, B.; Zhou, C.; Wang, L.; Song, C.; Jia, Y. Millimeter-Wave Radar-Based ECG Reconstruction Using Respiratory Harmonic Suppression and CA-WTBNet. *Bioengineering* **2026**, *13*, 731. [CrossRef] [PubMed]
62. Awal, M.A.; Mostafa, S.S.; Ahmad, M.; Rashid, M.A. An adaptive level dependent wavelet thresholding for ECG denoising. *Biocybern. Biomed. Eng.* **2014**, *34*, 238–249. [CrossRef]

63. Radev, H.; Petrova, G. An experimental setup for collecting physiological data from vehicle drivers. In Proceedings of the 15th International Scientific Conference TechSys 2026—Engineering, Technology and Systems, Plovdiv, Bulgaria, 14–16 May 2026. [CrossRef]
64. Sargolzaei, A.; Faez, K.; Sargolzaei, S. A new robust wavelet based algorithm for baseline wandering cancellation in ECG signals. In Proceedings of the IEEE International Conference on Signal and Image Processing Applications (ICSIPA), Kuala Lumpur, Malaysia, 18–19 November 2009. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.