# A Comparison of Wearable Tonometry, Photoplethysmography, and Electrocardiography for Cuffless Measurement of Blood Pressure in an Ambulatory Setting

Rebecca Mieloszyk, *Member, IEEE*, Hope Twede, Jonathan Lester, Jeremiah Wander, Sumit Basu, *Member, IEEE*, Gabe Cohn, *Member, IEEE*, Greg Smith, Dan Morris, Sidhant Gupta, Desney Tan, Nicolas Villar, Moni Wolf, Sailaja Malladi, Matt Mickelson, Lauren Ryan, Lindsey Kim, Jeffrey Kepple, Susanne Kirchner, Emma Wampler, Riena Terada, Joel Robinson, Ron Paulsen, and T. Scott Saponas, *Member, IEEE*

***Abstract*—*Objective:*** **While non-invasive, cuffless blood pressure (BP) measurement has demonstrated relevancy in controlled environments, ambulatory measurement is important for hypertension diagnosis and control. We present both in-lab and ambulatory BP estimation results from a diverse cohort of participants. *Methods:* Participants (N=1125, aged 21-85, 49.2% female, multiple hypertensive categories) had BP measured in-lab over a 24-hour period with a subset also receiving ambulatory measurements. Radial tonometry, photoplethysmography (PPG), electrocardiography (ECG), and accelerometry signals were collected simultaneously with auscultatory or oscillometric references for systolic (SBP) and diastolic blood pressure (DBP). Predictive models to estimate BP using a variety of sensor-based feature groups were evaluated against challenging baselines. *Results:* Despite limited availability, tonometry-derived features showed superior performance compared to other feature groups and baselines, yielding prediction errors of 0.32±9.8 mmHg SBP and 0.54±7.7 mmHg DBP in-lab, and 0.86±8.7 mmHg SBP and 0.75±5.9 mmHg DBP for 24-hour averages. SBP error standard deviation (SD) was reduced in normotensive (in-lab: 8.1 mmHg, 24-hr: 7.2 mmHg) and younger (in-lab: 7.8 mmHg, 24-hr: 6.7 mmHg) subpopulations. SBP SD was further reduced 15–20% when constrained to the calibration posture alone. *Conclusion:* Performance for normotensive and younger participants was superior to the general population across all feature groups. Reference type, posture relative to calibration, and controlled vs. ambulatory setting all impacted BP errors. *Significance:* Results highlight the need for demographically diverse populations and challenging evaluation settings for BP estimation studies. We present the first public dataset of ambulatory tonometry and cuffless BP over a 24-hour period to aid in future cardiovascular research.**

***Index Terms*—Cuffless blood pressure, hypertension, photoplethysmography, tonometry, wearable.**

Manuscript received August 12, 2021; revised December 23, 2021 and February 4, 2022; accepted February 5, 2022. Date of publication February 24, 2022; date of current version July 4, 2022. *(Corresponding author: Rebecca Mieloszyk.)*

Rebecca Mieloszyk, Hope Twede, Jonathan Lester, Jeremiah Wander, Sumit Basu, Gabe Cohn, Greg Smith, Dan Morris, Sidhant Gupta, Desney Tan, Nicolas Villar, Moni Wolf, Sailaja Malladi, Ron Paulsen, and T. Scott Saponas are with Microsoft Research, Redmond, WA 98052 USA (e-mail: becky@microsoft.com; hopetw@microsoft.com; jlester@microsoft.com; miah@microsoft.com; sumitb@microsoft.com; gabe@microsoft.com; gregsmi@microsoft.com; dan@microsoft.com; sidhant@microsoft.com; desney@microsoft.com; nvillar@microsoft.com; moniwolf@microsoft.com; samall@microsoft.com; ropau@microsoft.com; ssaponas@microsoft.com).

Matt Mickelson, Lauren Ryan, Riena Terada, and Joel Robinson are with Oculus VR, Redmond, WA 98052 USA (e-mail: mmickels@gmail.com; ldoneganryan@gmail.com; riena.terada@gmail.com; jkrobinso@msn.com).

Lindsey Kim and Jeffrey Kepple are with the Department of Medicine, Creighton University, Omaha, NE 68178 USA (e-mail: lindseyk1212@gmail.com; jeffkepple@live.com).

Susanne Kirchner is with Ideamed Digital and Kirinus Health, 80637 München, Germany (e-mail: susanne@adelhardt.co).

Emma Wampler is with Google, Seattle, WA 98103 USA (e-mail: ekwampler@gmail.com).

Digital Object Identifier 10.1109/JBHI.2022.3153259

## I. Introduction

HYPERTENSION, or high blood pressure, is one of the most significant risk factors for heart disease and stroke with an estimated prevalence of 47.3% in US adults over the age of 20 [1]. Despite the availability of numerous anti-hypertensive medications that are effective for the majority of patients, rates of uncontrolled hypertension remain high across the globe with a prevalence of 82.9% in an international cohort of hypertensive adults over 50 and 78.4% of US hypertensive adults [1], [2]. This is due in large part to challenges in how blood pressure is measured.

In practice, the clinical standard for the diagnosis of hypertension involves multiple blood pressure (BP) measurements in an office setting while the patient is at rest with confirmation over multiple office visits. The process can take months or even years to complete and is prone to multiple sources of error [3]. It is well established that ambulatory blood pressure monitoring (ABPM) is superior to office blood pressure measurement for both diagnosis and management of hypertension [4] as it can shorten the

time to diagnosis and improve diagnostic accuracy by assessing a patient's BP over the course of a 24-hour period. Furthermore, ABPM has the ability to disambiguate clinically meaningful hypertensive phenotypes such as masked and white-coat hypertension [5]. By measuring BP throughout the day and night, ABPM more comprehensively captures overall hypertensive load associated with downstream cardiovascular risk [6]–[8] and is correspondingly a better predictor of clinical outcomes [9], [10]. Despite these advantages, widespread clinical uptake of ABPM has yet to occur [3], [11]. Patient discomfort, sleep disturbance, and reimbursement issues are often cited as root causes for the low usage rates of ABPM in clinical practice [12], [13].

To counter these limitations, alternative approaches have been proposed to potentially overcome the shortcomings of brachial, cuff-based ABPM devices [14]–[25]. The majority of these are wearable devices designed to be worn on the wrist [14], [16], chest [17], or as a patch on the skin [18], [19]. There are also cuffless BP devices designed for continuous operation at the bedside [20]–[22] or ambient use [24]–[26]. Accuracy of these wearable devices has often been assessed in comparison to a reference oscillometric cuff [16]–[19], [25]; however, dual-observer auscultation is generally accepted as the gold-standard BP reference [27].

Cuffless BP devices allow for indirect BP measurement by leveraging hemodynamic phenomena that are known to change in conjunction with changes in BP [28], specifically pulse wave velocity and morphological characteristics of the pulsatile pressure or volume waveform in the arterial tree. These features are obtained via various sensing modalities: including radial tonometry, photoplethysmography (PPG), electrocardiography (ECG), and accelerometry. Tonometry requires applanation of the arterial lumen and subsequent transduction of the intra-arterial pulse pressure waveform. PPG is an optical sensing technique used to infer pulsatile volumetric changes most typically in distal vascular beds.

When joined with sensing of the electrical activity of the heart via ECG, tonometry or PPG can be used to measure the radial pulse arrival time (rPAT). Even for equivalent sensor positioning, though, optical rPAT is typically longer in duration than tonometric rPAT due to capillary diffusion time—the green wavelength used for reflective PPG senses activity in superficial vessels rather than the larger arteries whose pressure changes are detected by tonometry [28]. Pulse transit time (PTT) is equivalent to the sum of rPAT and the systolic timing interval known as the pre-ejection period (PEP). Ignoring the potentially substantial contribution of PEP, the rPAT timing interval is often employed as an inexact proxy for PTT and used to estimate BP [29]–[31] through a derivation of the Moens-Korteweg equation [32].

Periodic calibration is required for cuffless BP measurements that rely on estimation of change in BP with respect to a known reference. Calibration is performed using standard measurement techniques with an expected performance degradation as: (1) time in use exceeds the recommended calibration interval, and (2) conditions of use deviate from those of calibration [33]. In many cases, the algorithms used for estimation of BP employ supervised machine learning models trained on relatively small, experimentally convenient populations that may not generalize to the eventual user population. Taken jointly, these conditions suggest that the performance characteristics of a given cuffless BP device would degrade substantially when applied in real-world scenarios with clinically relevant patient populations.

TABLE I
COMPARISON OF PROTOCOL MEASUREMENTS

| Phase | Auscultatory Protocol | Oscillometric Protocol |
|---|---|---|
| Initial In-Lab Visit | Manual series: at rest/postural | Automated series: at rest/postural |
| $\geq$24 Hours Between Visits | *(no measurements)* | Automated ABPM, daytime every 30 mins nighttime every 60 mins |
| Return In-Lab Visit | Manual series: at rest/postural | Automated series: at rest/postural |

While some recent cuffless BP studies have recruited large, diverse cohorts [14], [15], many published cuffless BP device evaluations tend to measure predictive device performance in only young and healthy participant populations [27]. In this study, we sought to evaluate the performance of multiple cuffless BP measurement techniques in an ambulatory setting and on a diverse subject population that reflects the heterogeneity expected in clinical application. Specifically, we compare BP estimates derived from tonometric and optical sensing devices to BP measurement references including the gold-standard dual-observer auscultation and a clinically validated automated ocsillometric ABPM device. To our knowledge, this is the first large-scale study to assess performance of wearable tonometric and optical sensing devices on both healthy and hypertensive adults in an ambulatory setting.

## II. METHODS

### A. Study Protocol

Institutional review board (IRB) approval for this study was obtained from WCG IRB (Puyallup, WA, USA). Individuals unable to consent in English, pregnant women, prisoners, institutionalized individuals, and individuals younger than 18 were excluded from participation due to their vulnerable status. In support of the overall goal of assessing the clinical feasibility of multiple cuffless BP techniques, subject recruitment intentionally targeted a heterogeneous pool of individuals. Participants were selected to balance across gender and self-reported hypertensive status (see Tables II and III for resulting distributions). We saw this as critical to test the hypothesis that cuffless BP performance would differ across these populations.

Two separate protocols were performed—**auscultatory** and **oscillometric**—each using a different clinically-validated BP measurement technique. These protocols were non-concurrent, with each individual participant included in only a single protocol. In both protocols, a series of measurements were collected in two controlled in-lab visits at least 24 hours apart. During the oscillometric protocol, ambulatory measurements were also

TABLE II
STUDY DEMOGRAPHICS AND SUMMARY STATISTICS

| | Auscultatory | Oscillometric | Total |
|---|---|---|---|
| **Demographics** | | | |
| N analyzed | 642 | 483 | 1125 |
| Age (years) | 44.8 (12.8) | 45.6 (9.0) | 45.1 (11.3) |
| Female (%) | 49.7 | 48.7 | 49.2 |
| Height (in) | 67.3 (3.9) | 67.7 (4.0) | 67.5 (3.9) |
| Weight (lb) | 185.4 (49.0) | 198.2 (52.0) | 190.9 (50.7) |
| BMI (lb/in$^2$) | 28.6 (6.5) | 30.3 (7.2) | 29.3 (6.9) |
| History of hypertension (%) | 29.6 | 40.6 | 34.3 |
| History of other CVD (%) | 11.4 | 14.9 | 12.9 |
| **Observed Blood Pressure** | | | |
| Supine SBP (mmHg) [N=642, 479] | 128 (18) | 126 (15) | |
| Supine DBP (mmHg) | 78 (12) | 78 (10) | |
| Seated SBP (mmHg) [N=452, 479] | 125 (19) | 130 (15) | |
| Seated DBP (mmHg) | 76 (12) | 86 (10) | |
| 24-hr SBP (mmHg) [N=483] | | 124 (14) | |
| 24-hr DBP (mmHg) | | 76 (9) | |

Parenthetical values indicate standard deviation.

TABLE III
PARTICIPANT SELF-REPORTED HYPERTENSIVE PROFILE

| Age | Normal or Unknown Status | Controlled Hypertension | Uncontrolled Hypertension | Not Recorded |
|---|---|---|---|---|
| 21-29 | 69 | 3 | 6 | 1 |
| 30-39 | 270 | 54 | 42 | 1 |
| 40-49 | 195 | 79 | 44 | 3 |
| 50-59 | 165 | 109 | 78 | |
| 60+ | 45 | 49 | 23 | |

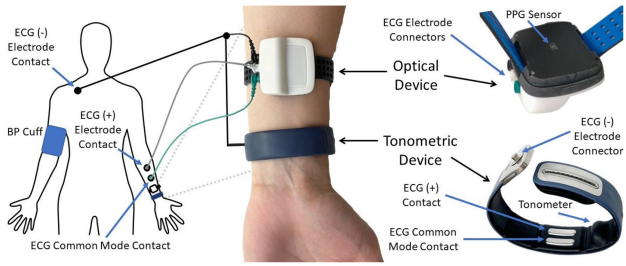



Fig. 1. Experimental setup showing tonometric and optical sensing devices devices (*right*), and placement of devices, BP cuff, and ECG connections (*left*).

collected during the time between the in-lab visits. Ambulatory measurements were not feasible in the auscultatory protocol. Table I summarizes both protocols.

*1) Participant Demographics:* In total, 1315 participants were enrolled in the study across both protocols. Of those, 1125 (85.6%) completed their respective protocol with sufficient usable data to be included in subsequent analyses. There were numerous reasons why participants' data were not available for subsequent use, but the most common causes were discontinuation of the protocol due to device discomfort, unwillingness to participate, or technical issues with one or more measurement devices. Demographic characteristics and summary statistics of the study population included in subsequent analyses are shown in Table II. It is interesting, though unsurprising, to note that there was a discrepancy between self-reported hypertensive status shown in in Table III and observed resting seated blood pressures. For the 553 subjects from both protocols with successful in-lab seated BP measurements who self-reported normotensive status, 274 (49.5%) would have been classified as having hypertension per the 2017 AHA guidelines [34]. Note the putative association between a participant's age and their self-reported hypertensive status shown Table III.

*2) Device Placement:* In both protocols, participants were fitted with three devices (see Fig. 1): a cuff-based, clinically-validated BP measurement device, a tonometric sensing device, and an optical sensing device. The tonometric device was worn on the wrist and fitted with a pressure sensor to measure the pulse pressure waveform at the radial artery. An extender could be used to adjust for participant comfort or signal strength of the tonometer. The optical sensing device was worn proximal to the tonometric sensing device. The optical sensing device used a standard PPG sensor placed on the anterior surface of the arm. Additionally, both devices had an ECG sensor (ADS1292) used to measure cardiac electrical activity and to time-align measurements across devices, with electrodes applied to measure Lead I ECG. The BP cuff was worn on the opposite arm of the tonometric and optical devices. BP cuff size was determined by measuring the upper arm circumference and selecting a cuff according to manufacturer specifications.

*3) Tonometric Sensing Device:* Radial tonometric sensing is typically performed by a trained operator who holds a stylus equipped with a pressure sensor tip on the patient's radial artery. The operator must be skilled at finding the patient's radial artery and applanating the artery without deviating from or constricting the vessel, which would result in a poor measurement signal [35]–[37]. Wearable tonometric devices have been designed [38] to allow more continuous measurement of radial pulse pressure waveforms in clinical settings. An earlier version of the tonometric device presented in this manuscript has been validated for its ability to measure clinically relevant morphological characteristics of the radial pressure wave [39].

The wrist-worn design of the tonometric device enables non-experts to properly place the sensor and capture high quality radial pulse pressure waves. The electronics in the tonometric device are distributed around the wrist to improve comfort with long-term wear, and a force concentrator is used over the pressure sensor (MS5803-02BA) to maximize signal quality. ECG sensing is provided via two exposed stainless steel bumps on the inner wrist (dry contacts) and a connection point for a standard wet electrode. The wet electrode wire is run up the participant's arm and connected to the chest on the opposite side of the heart (see Fig. 1). An accelerometer (LSM6DS3) is used to monitor arm posture and activity level.

*4) Optical Sensing Device:* The optical device resembles a large smart watch and records ECG and PPG signals. It incorporates a commercial PPG optical sensor (MAX30101) that sits underneath the watch body. The three ECG contacts are made via standard wet electrodes with two electrodes connected on the same arm as the device, and the third connected to the chest on the opposite side of the heart (see Fig. 1). An accelerometer (LSM330) is used to monitor arm posture and activity level.

*5) Auscultatory Protocol:* The auscultatory protocol consisted of two in-lab visits where a series of dual-observer auscultation BP measurements were taken by trained observers using the ADC Diagnostix 700/703 aneroid sphygmomanometer. The brachial cuff was fitted to the participant's left or right arm (randomized). Cuff inflation and deflation were manually controlled by one of the observers. Measurements were taken in supine and seated postures.

*6) Oscillometric Protocol:* In the oscillometric protocol, BP was measured using the Spacelabs Healthcare OnTrak 90227 ambulatory blood pressure monitor. The brachial cuff was fitted to the participant's dominant arm or according to the participant's preference. Cuff inflation and deflation were automatically controlled by the ABPM device. During the in-lab phases, measurements were manually initiated by the observers and taken in both supine and seated postures. During the ambulatory phase, measurements were triggered automatically every 30 minutes during approximate waking hours (8AM–8PM) and every 60 minutes at nighttime (8PM–8AM). These measurement intervals are consistent with recommendations from the European Society of Hypertension [40]. The return visit consisted of participant feedback and a second series of automated oscillometric measurements. The participant was asked whether they removed any of the devices during the 24-hour ambulatory phase, and if so, they were asked approximately when and for how long. Further protocol details accompany our Aurora-BP study data and are accessible at https://github.com/aurorabp/sample-data.

## B. Data Collection and Preprocessing

The method for collection of data from the tonometric and optical sensing devices varied based the phase of the protocol that was being executed. During the in-lab visits, data were streamed from the devices in real time via Bluetooth using custom software written in C#. The devices streamed continuously for the duration of the in-lab visit, and time indices were recorded by the observers to denote protocol progression and timing of reference measurements. Post-hoc manual review of the time indices and data streams was used to ensure that the experimental protocol was followed and that the devices were functioning correctly; otherwise, data were excluded from subsequent analyses. The raw data from in-lab phases were then segmented into measurement windows based on marked time indices. All subsequent operations, including feature extraction, were performed on these per-measurement data windows. During automated measurements taken by the oscillometric device, windows were standardized to 30 seconds in duration. Timing of the auscultatory process was variable, and thus the measurement window duration correspondingly varied (median 20.1 sec, IQR 4.9 sec).

During ambulatory wear, data for each measurement were stored locally on the devices and retrieved at a later time using custom software written in C#. Precise control of scheduled measurement timing for the oscillometric ABPM device was not possible; thus, the tonometric and optical sensing devices were configured to record data 2 min before and 2 min after the expected timing of the oscillometric reference measurement. Oscillometric measurement timing was then taken from the measurement logs of the oscillometric ABPM device and a 15 sec window was segmented out for subsequent analysis. Synchronization with the oscillometric ABPM device during ambulatory measurements was limited by this approach, as it was only accurate to within the limits of clock synchronization and drift.

*1) Filter Details:* Tonometry, PPG, and ECG signals were then filtered to reduce noise and isolate spectral content of interest. All signal preprocessing was conducted using custom software developed in Matlab. For the ECG signal, the following filters were implemented: a DC block filter (0.1 Hz cutoff), lowpass 7th-order elliptical filter (40 Hz passband, 45 Hz stopband, 0.1 dB passband ripple), and a notch 6th-order Chebyshev type I filter (60 Hz center, 3 Q-factor, 0.1 dB passband ripple). For the PPG signal, the following filters were implemented: a highpass 4th-order Butterworth filter (0.25 Hz cutoff) and a lowpass equiripple filter (10 Hz passband, 12 Hz stopband, 1 dB passband ripple, 60 dB stopband attenuation). For the tonometric signal, the following filters were implemented: a highpass elliptical filter (0.2 Hz stopband, 0.3 Hz passband, 60 dB stopband attenuation, 1 dB passband ripple) and a lowpass 7th-order elliptical filter (26 Hz stopband, 22 Hz passband, 0.1 dB passband ripple).

*2) Time Alignment and Segmentation:* The sensor data from a single device are inherently time-aligned; however, clock drift and other factors necessitate time alignment between different sensing devices. To perform this alignment, we identified QRS complexes in the ECG signals from each device in a measurement window using a modified Pan-Tompkins algorithm [41], constructed derived signals that contained only the ECG peaks, and time-shifted the data from one device such that the maximal cross-correlation occurred at zero lag. These alignments were

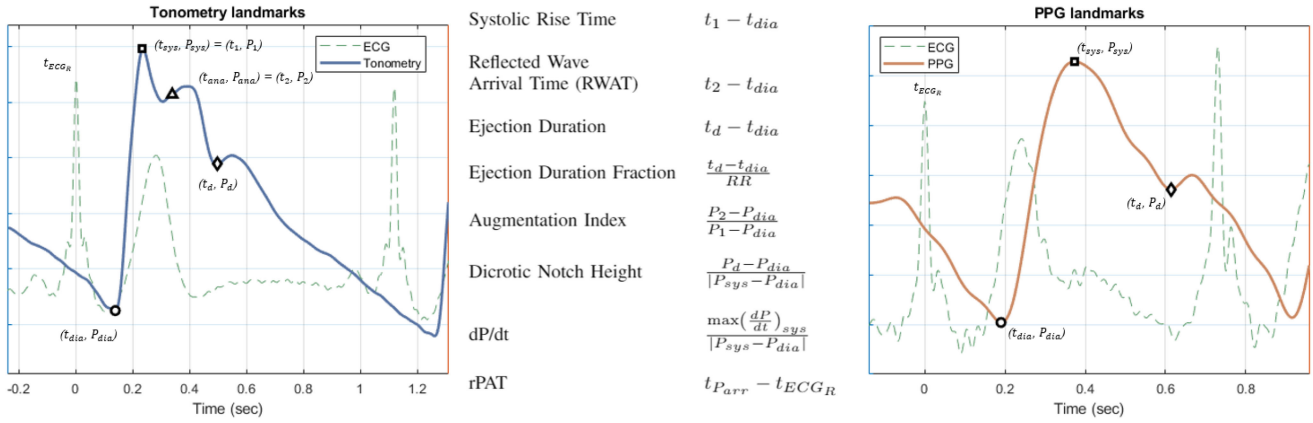



Fig. 2. Example tonometric (*left*) and PPG (*right*) waveforms time-aligned with ECG. Note the presence of anacrotic and dicrotic notches in the tonometric waveform and dicrotic in the PPG waveform. The forward wave pressure, $P_1$, refers to the anacrotic notch pressure ($P_{ana}$) for an $a$-type wave and systolic pressure ($P_{sys}$) for a $c$-type wave as shown here. The reflected wave pressure, $P_2$, then refers to $P_{sys}$ in an $a$-type wave and $P_{ana}$ in a $c$-type wave [43].

automatically performed, but manually reviewed, and windows with failed alignments were removed from subsequent analyses. We then performed pulse segmentation for the tonometry and PPG waveforms, again leveraging a modified Pan-Tompkins algorithm.

### C. Signal Quality

We computed a heuristic-based quality score that penalized measurement data windows with substantial artifacts, poor signal-to-noise ratio, and lack of consistency between the pulses within a window. This signal quality metric was indicative of confidence that features derived from a given measurement were truly physiologic in nature and was thresholded to determine which measurements were excluded from subsequent training and testing of predictive models.

### D. Feature Extraction

In order to estimate BP and compare with the references measured via auscultation and oscillometry, features were extracted from tonometric, PPG, ECG, and accelerometer waveforms. Given the known relationship between rPAT, PTT, and blood pressure [28], we included rPAT-related features as well as other morphological waveform characteristics [42]. Morphological feature values were computed from each available pulse and then aggregated by taking the median over the all pulses in the measurement window. All other features were computed across the entire measurement window. In addition to sensor-based features, the cosine and sine of the hour of day (normalized by $2\pi/24$) were also included in all feature groups in order to help the models compensate for diurnal changes in BP.

Table IV summarizes the feature groups, labeled: "ECG," "optical," "smart watch," and "tonometric". Note that many features can be derived using different sensors as shown in Fig. 2. For example, there is both an optical rPAT feature obtained from the PPG sensor waveform and a separate tonometric rPAT feature obtained from the tonometric sensor waveform.

TABLE IV
FEATURES

| Feature | Feature Group | | | |
|---|---|---|---|---|
| | ECG | Optical | Smart watch | Tonometric |
| Arm Angle | · | ✓ | ✓ | ✓ |
| Augmentation Index | · | · | · | ✓ |
| Beat Length | ✓ | ✓ | · | ✓ |
| Dicrotic Notch Height | · | · | · | ✓ |
| dP/dt | | | | |
| *Optical* | · | ✓ | ✓ | · |
| *Tonometric* | · | · | · | ✓ |
| Ejection Duration | · | · | · | ✓ |
| Ejection Duration Frac. | · | · | · | ✓ |
| Heart Rate | | | | |
| *ECG* | ✓ | ✓ | · | ✓ |
| *Optical* | · | ✓ | ✓ | · |
| *Tonometric* | · | · | · | ✓ |
| Heart Rate Variability | | | | |
| *ECG* | ✓ | ✓ | · | ✓ |
| *Optical* | · | · | ✓ | · |
| Quality | | | | |
| *Optical* | · | ✓ | ✓ | · |
| *Tonometric* | · | · | · | ✓ |
| Inverse rPAT | | | | |
| *Optical* | · | ✓ | · | · |
| *Tonometric* | · | · | · | ✓ |
| rPAT | | | | |
| *Optical* | · | ✓ | · | · |
| *Tonometric* | · | · | · | ✓ |
| RWAT | · | · | · | ✓ |
| Systolic Rise Time | · | · | · | ✓ |
| Time of Day | | | | |
| $cos\left(2\pi\frac{hour}{24}\right)$ | ✓ | ✓ | ✓ | ✓ |
| $sin\left(2\pi\frac{hour}{24}\right)$ | ✓ | ✓ | ✓ | ✓ |

The tonometric feature group contains mostly tonometry-derived features, and the optical feature group similarly contains mostly features directly derived from the PPG waveform. The ECG, optical, and tonometric feature groups also include features related to heart rate derived from ECG, while the smart watch feature group is intended to provide a baseline akin to a fitness wearable equipped with only PPG and accelerometer sensing.

### E. Predictive Modeling

We employ machine learning models to predict blood pressure and use the model performance to evaluate the differences among feature groups, subpopulations, and BP reference methods. Separate models were trained using each of the feature groups shown in Table IV and using different subpopulations. In addition to evaluating on the data from all study participants, we analyzed subsets of the data from younger (age $\leq$35) and normotensive (resting SBP $\leq$130 mmHg) subpopulations commonly studied the BP literature [23], [25], [26], [44].

To predict the change in BP from the in-lab calibration values measured during the initial visit, we evaluated a range of machine learning methods with substantially similar performance when trained on the features listed in Table IV. These methods included generalized additive models, gradient-boosted regression trees, ridge regression, an approach we refer to as variance-optimized regression (which minimized a combined loss of mean squared error, bias, and standard deviation), and literature-based BP prediction models [33], [44]–[46]. Ridge regression [47] was found to perform as well as or better than all of the alternate models. Due to its parsimony, interpretability, and high performance, only ridge regression results are presented.

Prediction results were also compared to a number of challenging baselines. The "zero baseline" assumes no change in BP from the time of initial calibration. The "static baseline" uses the in-lab initial visit calibration values (SBP and DBP) and the time of day features listed in Table IV as inputs to a ridge regression model, allowing a modicum of flexibility in estimation, while not leveraging any features derived from the tonometric or optical sensing devices.

Cross-validation was performed with folds stratified by participant—all measurements associated with a particular individual were held out when predicting that individual's data. This stratification strategy is critical to prevent data leakage and provide realistic estimates of performance on unseen individuals. In order to estimate confidence intervals and other statistics on these quantities, we repeated the cross-validation procedure on 100 bootstrapped [48] samples of the participant pool, with each bootstrap replicate containing 90% of the participants. This resulted in a population of 100 values for each bias and standard deviation quantity, which was used to produce confidence intervals and the reported mean values. To compute this statistical significance between the mean values, we used Welch's unequal variances t-test [49], which is meant for comparing two sample populations drawn from distributions of unequal variances.

For in-lab BP measurements, we predicted the return visit BP taken at rest at least 24 hours after the initial in-lab calibration. In-lab return visit predictions were compared to simultaneously collected reference BP measurements using clinically validated devices (auscultatory or oscillometric, depending on the protocol) in order to obtain the commonly reported bias and standard deviation measures of cuffless BP devices with respect to a reference [27].

For ambulatory BP measurements, a prediction was made for each reference measurement taken by the oscillometric device during the $>$24 hours of wear. These predictions were then aggregated into a single 24-hour average using the following method. To account for different sampling periods during day and night, predictions were binned and averaged by hour, producing a single individual hour-average value for each hour of the day. A simple mean of these 24 values then resulted in the single 24-hour average. The individual hour-average values were compared to the 24-hour average values similarly computed from the automated oscillometric reference device, resulting in a signed error value for each individual. Errors were then aggregated into bias and standard deviation values, which were then computed for each bootstrap replicate across all individuals.

## III. Results

### A. Blood Pressure Variability

Since the majority of cuffless blood pressure devices base their measurements on an initial calibration, blood pressure variability (BPV) is a highly relevant consideration when assessing a device's performance. BPV was markedly different among the various age groups and hypertensive profiles represented in the data. Fig. 3 uses the absolute change in systolic blood pressure ($\Delta$SBP) as a proxy for BPV and shows an increasing degree of individual BPV observed in older and hypertensive groups, where normotensive is defined as resting SBP $<$ 130 mmHg.

The measurement environment also contributes to observed BPV. Note in Fig. 4 that the in-lab $\Delta$SBP standard deviation is lower than the ambulatory $\Delta$SBP standard deviation for all analyzed subgroups likely due to the greater range in BPV during daily activity captured during the ambulatory phase as compared to the resting poses of the in-lab phase. The measurement device also plays a role—overall, $\Delta$SBP is observed to be most tightly bound for auscultatory measurements. This would suggest noise contributions from the oscillometric measurement technique even when applied in controlled environments such as the in-lab setting.

### B. Data Yield and Signal Quality

*1) In-Lab Measurements:* During the in-lab phases of the protocols, trained observers were able to ensure measurement compliance and optimize signal quality, and thus the rate of obtaining suitable measurements for downstream analyses was extremely high. A total of 88.1% of participants from the oscillometric protocol and 93.3% from the auscultatory protocol (from 548 and 688 total subjects, respectively) yielded in-lab measurements usable for downstream analyses.

*2) Ambulatory Measurements:* For any medical device that is intended to be used outside of a controlled environment or clinical setting, it is important to understand whether users of the device will be capable of and willing to use the device as instructed. Though a detailed breakdown of the reasons for missed or invalid ambulatory measurements is beyond the scope of this manuscript, we provide a brief assessment below as it pertains to subsequent analyses.

The precise duration of wear for the ambulatory phase of the study varied based on appointment times for initial and return visits. As participants were instructed to wear the measurement

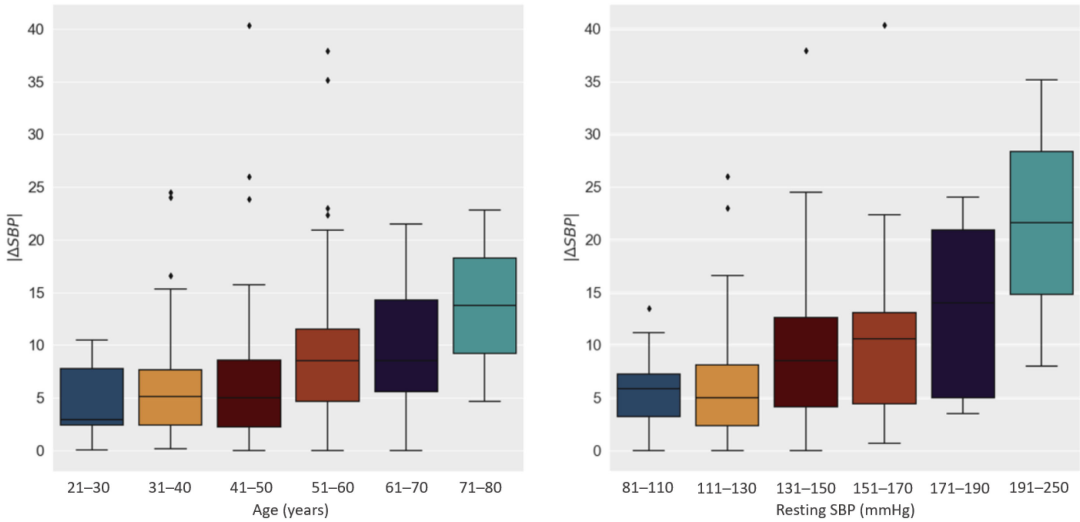



Fig. 3. Change in auscultatory SBP measurement over 24 hours from initial to return visit increases with both age (*left*) and resting SBP (*right*). Boxes are bound by first and third quartiles of the data in each bin. Participants over 50 years of age and those approaching hypertensive resting SBP ranges greater than 130 mmHg exhibit notably different mean absolute SBP change from initial to return visit.

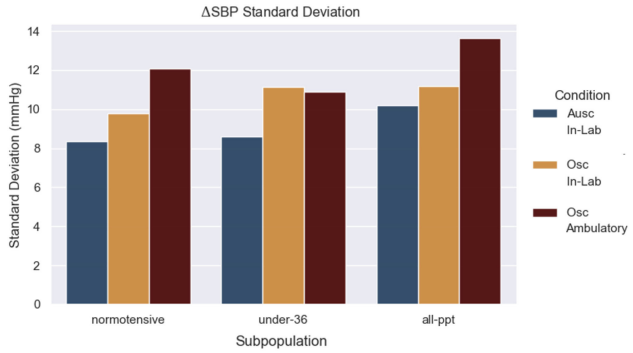



Fig. 4. Standard deviation of ΔSBP measurement by population subgroup (normotensive, age $\leq$ 35 years,, and all participants) for auscultatory in-lab, oscillometric in-lab, and oscillometric ambulatory measurements. Oscillometric conditions were most variable, while auscultatory measurements exhibited the tightest distribution for each subpopulation.

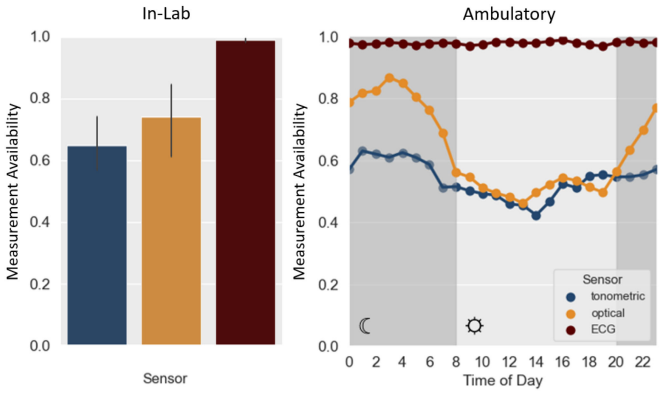



Fig. 5. Measurement availability in controlled in-lab (*left*) and ambulatory settings (by time of day, *right*). ECG measurements are consistently available, while both tonometric and optical availability are increased during nighttime wear.

apparatus for the duration of the time between these visits, the number of ambulatory measurements obtained varied correspondingly. The average number of attempted ambulatory measurements for each participant was 41.6 ($\sigma = 4.0$). Of these, 19.3 ($\sigma = 10.1$) usable ambulatory measurements were obtained per participant. The difference between these two figures includes missed measurements due to apparatus component removal, device malfunction, procedural non-compliance, and poor signal quality.

Participation in the study was voluntary, and the full measurement apparatus was both complicated and burdensome. According to participant-provided feedback of the 548 participants who completed the ambulatory phase of the oscillometric protocol, 31% of participants reportedly lost at least one measurement due to apparatus miswear or malfunction of one or more devices. More than 28% of the participants who completed the ambulatory phase self-reported removal of one or more component of the apparatus due to device malfunction or of the participant's volition, resulting in missed measurements. Of participants who reported removal of one or more component of the apparatus, over 52% did not replace the apparatus after initial removal and, notably, 77% reported apparatus removal at night.

Participants reported occasionally being unable or unwilling to follow measurement procedure, which also resulted in a missed ambulatory measurement. Even when devices were being worn and measurement protocols were being followed, some ambulatory measurements were unusable because the signal quality was insufficient for subsequent analysis. These two scenarios cannot be effectively disambiguated from the data and were assessed jointly.

*3) Measurement Availability by Sensor:* The measurement availability by sensor type is displayed in Fig. 5 for both in-lab and ambulatory phases. An optical or tonometric measurement is considered "available" when it exceeds the determined threshold

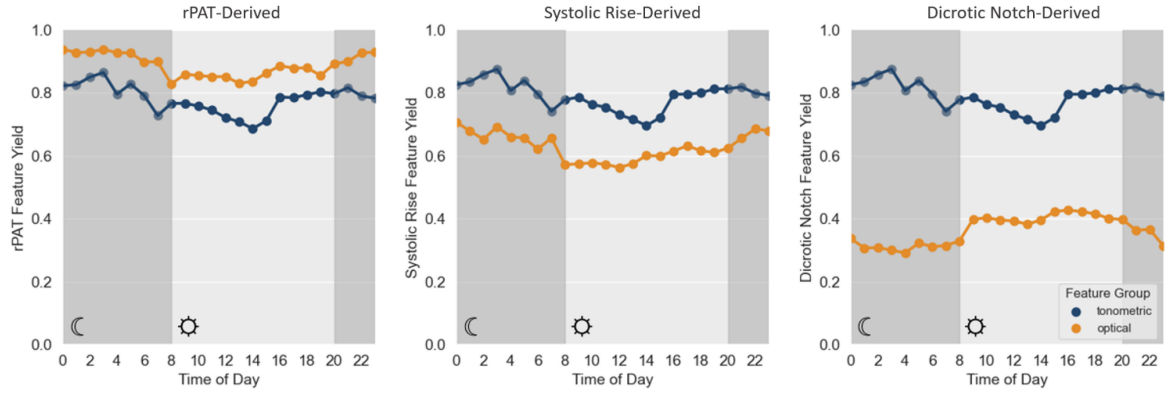
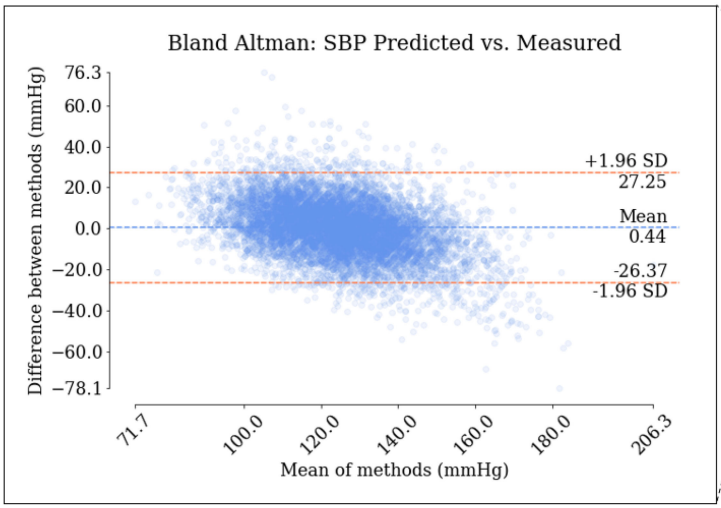



Fig. 6. Tonometric and optical feature yield for rPAT-derived features including {rPAT, inverse rPAT}, systolic rise-derived features including {dP/dt, systolic rise time}, and dicrotic notch-derived features including {dicrotic notch height, ejection duration, and ejection duration fraction}. Tonometric features exhibit higher yield than optical for features derived from systolic rise and dicrotic notch, but optical outperforms tonometric feature yield for rPAT-derived features.




24-hr Average Ambulatory BP Measurement Prediction

| Subpopulation and Feature Group | Oscillometric SBP Bias | Oscillometric SBP SD | Oscillometric DBP Bias | Oscillometric DBP SD |
|---|---|---|---|---|
| **All** | | | | |
| Tonometric | 0.855 | **8.66*** | 0.752 | 5.94 |
| Optical | **0.361*** | 8.94 | **0.430*** | 6.01 |
| ECG | 1.08 | 8.96 | 1.32 | 5.94 |
| Smart watch | 0.421 | 8.98 | 0.583 | 5.95 |
| Static Baseline | 1.10 | 8.93 | 1.34 | **5.85*** |
| Zero Baseline | 3.61 | 10.2 | 2.45 | 6.65 |
| **Normotensive** | | | | |
| Tonometric | 0.964 | **7.23*** | 0.885 | **5.18*** |
| Optical | 0.556 | 7.45 | **0.634*** | 5.41 |
| ECG | 1.13 | 7.53 | 1.42 | 5.36 |
| Smart watch | 0.592 | 7.50 | 0.784 | 5.35 |
| Static Baseline | 1.12 | 7.60 | 1.43 | 5.31 |
| Zero Baseline | **0.491*** | 8.27 | 0.948 | 5.95 |
| **Age $\leq$ 35** | | | | |
| Tonometric | 2.22 | 6.71 | 2.34 | 4.55 |
| Optical | **1.39** | 6.52 | **1.65** | 4.40 |
| ECG | 1.87 | **6.15*** | 2.41 | **4.31*** |
| Smart watch | 1.43 | 6.48 | 1.68 | 4.50 |
| Static Baseline | 1.55 | 6.57 | 2.12 | 4.64 |
| Zero Baseline | 4.01 | 7.79 | 3.00 | 5.85 |

Fig. 7. 24-hr Average Ambulatory BP Measurement Prediction: Results represent 24-hr averages from participants with a minimum of 17 usable ambulatory measurements. Bolded values indicate the best performance for each column within each subpopulation (all, age $\leq$ 35 years, normotensive). Asterisks indicate significance with respect to Welch's unequal variances t-test at a $p = 0.05$ threshold (based on 100 bootstraps) relative to the nearest values in the column and subpopulation. The Bland Altman plot is shown for tonometric SBP prediction with 95% limits of agreement.

of the signal quality score described in Section II. For ECG, we define measurement availability as successful alignment with corresponding optical and tonometric signals. Measurements from the ECG sensor were most available, followed by optical and tonometric sensors. ECG measurement availability remained steady across time, while tonometric and optical measurement availability peaked during nighttime hours.

*4) Feature Yield:* Features were derived from the waveform landmarks annotated in Fig. 2. If the required landmarks for a given feature, such as the dicrotic notch landmark for the "Dicrotic Notch Height" feature, were undetected in a given window, that feature was not computed for that measurement. Fig. 6 compares feature yield for features dependent on detection of rPAT, systolic rise, and dicrotic notch. The optical rPAT-derived feature yield was slightly greater than tonometric, but tonometric features related to the dicrotic notch and systolic rise exhibited greater yield than optical. This was most notable for dicrotic notch features, which were more than twice as available for the tonometric feature group. We also show the results of binning feature yield by hour of day. Both tonometric and optical features are slightly more available during nighttime hours than daytime, when participants tended to be moving more. Measurements and extracted features are available at https://github.com/aurorabp/sample-data.

## C. Modeling

*1) Comparison of Participant Subsets:* Both in-lab return visit (Table V) and 24-hour average ambulatory (Fig. 7) BP prediction errors are reported as mean bias and standard deviation

TABLE V
RETURN VISIT BP MEASUREMENT PREDICTION

| Subpopulation and Feature Group | Auscultatory | | | | Oscillometric | | | |
|---|---|---|---|---|---|---|---|---|
| | SBP | | DBP | | SBP | | DBP | |
| | Bias | SD | Bias | SD | Bias | SD | Bias | SD |
| **All** | | | | | | | | |
| Tonometric | **0.316*** | **9.75*** | **0.542** | 7.65 | −1.12 | 11.3 | **−0.006 84*** | 8.18 |
| Optical | 0.434 | 10.0 | 0.562 | **7.60*** | −0.684 | 11.1 | −0.341 | 8.11 |
| ECG | 0.377 | 10.2 | 0.591 | 7.82 | **−0.592*** | 11.1 | 0.362 | 8.29 |
| Smart watch | 0.417 | 10.1 | 0.548 | 7.66 | −0.759 | **11.0** | −0.315 | 8.24 |
| Static Baseline | 0.387 | 10.2 | 0.616 | 7.87 | −0.691 | **11.0** | −0.182 | **7.96*** |
| Zero Baseline | 1.72 | 10.3 | 1.62 | 8.12 | −5.79 | 11.5 | −7.96 | 8.19 |
| **Normotensive** | | | | | | | | |
| Tonometric | 0.0580 | **8.06*** | **0.439*** | **7.14*** | −1.51 | 10.2 | −0.361 | 7.91 |
| Optical | 0.121 | 8.40 | 0.512 | 7.28 | −0.659 | 10.0 | −0.334 | 7.93 |
| ECG | **−0.009 01** | 8.53 | 0.537 | 7.38 | **−0.307*** | 10.1 | 0.533 | 7.98 |
| Smart watch | 0.0666 | 8.47 | 0.503 | 7.36 | −0.556 | 10.0 | −0.135 | 8.17 |
| Static Baseline | −0.0276 | 8.55 | 0.557 | 7.40 | −0.438 | **9.97** | **−0.0544*** | **7.57*** |
| Zero Baseline | 0.431 | 8.60 | 0.851 | 7.69 | −7.62 | 10.2 | −8.74 | 7.79 |
| **Age $\leq$ 35** | | | | | | | | |
| Tonometric | 0.376 | **7.77*** | **0.608*** | **7.81*** | **0.172*** | 11.8 | 1.20 | 8.81 |
| Optical | 0.342 | 7.95 | 0.713 | 7.86 | 0.818 | **10.1*** | 0.503 | 8.06 |
| ECG | 0.386 | 8.23 | 0.939 | 8.08 | 1.20 | 11.2 | 0.732 | 8.27 |
| Smart watch | **0.318** | 8.02 | 0.789 | 8.03 | 0.902 | 10.3 | 0.474 | 8.29 |
| Static Baseline | 0.467 | 8.18 | 0.978 | 8.07 | 0.912 | 11.1 | **0.0751*** | **7.60*** |
| Zero Baseline | 2.38 | 8.20 | 1.52 | 8.50 | −4.57 | 11.8 | −8.34 | 8.52 |

Bolded values indicate the best performance for each column within each subpopulation (all, normotensive, age$\leq$35). Asterisks indicate significance with respect to Welch's unequal variances t-test at a $p = 0.05$ threshold (based on 100 bootstraps) relative to the nearest values in the column and subpopulation.

TABLE VI
MODEL COMPARISON FOR TONOMETRIC 24-HR

| Model Type | SBP | | DBP | |
|---|---|---|---|---|
| | Bias | SD | Bias | SD |
| Generalized Additive Models | 0.758 | 8.76 | 0.417 | 6.01 |
| Gradient-Boosted Decision Trees | 0.805 | 8.86 | 0.497 | 6.04 |
| Ridge Regression | 0.823 | 8.56 | 0.558 | 5.93 |
| Variance-Optimized Regression | 0.691 | 8.84 | 0.554 | 6.02 |

over the 100 bootstrap replicates. The best-performing feature groups for each metric (columns) are bolded, and asterisks indicate where the best-performing value exhibits a significant difference relative to the second best. Here, statistical significance is defined with respect to Welch's unequal variances t-test at a $p = 0.05$ threshold.

Due to the nature of ambulatory monitoring, each participant had a different number of usable ambulatory measurements. Fig. 7 presents data from 227 participants with at least 17 usable measurements. See the shared Aurora-BP study for bias and standard deviation (as well as confidence intervals on these values) across a range of minimum number of usable ambulatory measurements. Especially for the tonometric feature group, standard deviation improved with an increasing number of minimum usable ambulatory measurements. However, the number of subjects retained for analysis correspondingly decreased. The 95% confidence intervals computed for both bias and standard deviation were tightly bound. Table VI presents a comparison of multiple model types described in Section II over 100 bootstrap samples for 24-hour prediction. As noted, these model types perform comparably for both SBP and DBP.

In a gross comparison, measurements from the all-participant population in Table V and Fig. 7 had an overall higher standard deviation than normotensive or young subsets. This indicates that, if analyzing the complete population, including older and hypertensive participants, performance of tonometric, optical, and ECG feature groups is worse than if only analyzing normotensive or young participants. This is notable within the context of other cuffless BP device studies, which often only analyze a small cohort of young, healthy participants.

Results are all referenced to a supine calibration. In order to examine posture-specific differences, Table VII reports standard deviation of auscultatory SBP prediction for different postures following supine calibration. As expected, the standard deviation is smallest in the supine posture since this matches the calibration. As in Table VII, we again see an improvement within normotensive as well as young subpopulations, which tend to have comparatively lower variation in blood pressure. Note that the combined effect of these factors is substantial—the supine posture results for normotensive and younger subpopulations perform well enough to meet criterion I of the ANSI/AAMI specification (standard deviation $<$8 mmHg, bias $<$5 mmHg) [27] for all feature groups and even all baselines. As such, this warrants a cautionary note—if a study does not ensure that its participant population has a broad distribution not only in age but also hypertensive status, they risk having results that are overly optimistic and may not generalize to the target population.

*2) Auscultation Vs. Oscillometric References:* Return visit prediction performance from the auscultatory protocol exceeds performance from the oscillometric protocol for nearly every condition shown in Table V. For all feature groups and subpopulations, SBP standard deviation is $\geq$1 mmHg higher for

TABLE VII
RETURN VISIT POSTURE VARIATION

| Subpopulation and Feature Group | Auscultatory SBP | | |
|---|---|---|---|
| | All | Supine | Seated |
| | SD | SD | SD |
| **All** | | | |
| Tonometric | **9.75*** | **9.27*** | **10.5*** |
| Optical | 10.0 | 9.60 | **10.5*** |
| ECG | 10.2 | 9.61 | 10.9 |
| Smart watch | 10.1 | 9.54 | 10.6 |
| Static Baseline | 10.2 | 9.62 | 10.9 |
| Zero Baseline | 10.3 | 9.70 | 10.9 |
| **Normotensive** | | | |
| Tonometric | **8.06*** | **7.39*** | **8.95*** |
| Optical | 8.40 | 7.71 | 9.08 |
| ECG | 8.53 | 7.69 | 9.38 |
| Smart watch | 8.47 | 7.64 | 9.24 |
| Static Baseline | 8.55 | 7.74 | 9.35 |
| Zero Baseline | 8.60 | 7.81 | 9.38 |
| **Age $\leq$ 35** | | | |
| Tonometric | **7.77*** | **7.00*** | 8.61 |
| Optical | 7.95 | 7.30 | **8.35*** |
| ECG | 8.23 | 7.28 | 8.83 |
| Smart watch | 8.02 | 7.21 | 8.48 |
| Static Baseline | 8.18 | 7.24 | 8.78 |
| Zero Baseline | 8.20 | 7.29 | 8.77 |

Standard deviation of auscultatory return visit SBP prediction following supine calibration during the initial visit. Bolded values indicate the best performance for each column within each subpopulation (all, normotensive, age$\leq$35). Asterisks indicate significance with respect to Welch's unequal variances t-test at a $p = 0.05$ threshold (based on 100 bootstraps) relative to the nearest values in the column and subpopulation.

oscillometry, even for the zero baseline. This may be indicative of a greater degree of inherent measurement noise in the oscillometric device as a reference as compared to manual auscultation.

*3) 24-Hour Average Vs. Single-Timepoint BP:* The overall 24-hour average prediction performance shows reduced standard deviation in comparison with oscillometric single-timepoint return visit measurements. This is to be expected if the errors are to some degree uncorrelated, as averaging over multiple measurements will reduce errors in this case; however, the observed reduction is far less dramatic than would be the case if the errors were independently distributed. Results show even further reduced standard deviation for normotensive and young participants. We further examined BP prediction error by time of day with details available in the shared Aurora-BP study. No substantial change in standard deviation was observed with time, and a slight variation in bias was observed as the models did not fully capture diurnal variation.

*4) Comparison of Feature Groups:* In general, tonometry outperformed the other feature groups for the most challenging metric: SBP standard deviation. The tonometric feature group's SBP standard deviation is significantly lower than the other feature groups and both baselines, with the exception of the young subset where all feature groups and baselines show reduced standard deviation.

DBP presents a much more consistent target in that its inherent variation is far less than SBP. As a consequence, DBP prediction bias and standard deviation values are both lower and much more similar across the feature groups and population subsets when compared to SBP.

*5) Performance Against Baselines:* The zero baseline exhibits the smallest bias for the normotensive group, indicating the consistency and comparatively low variability of BP values in this subpopulation. Conversely, the zero baseline stands out with the poorest performance in both SBP and DBP prediction for the entire population. As expected, the more flexible static baseline exhibits smaller standard deviation and much less bias than the zero baseline in nearly all conditions.

Notably, the optical feature group provides only marginal performance benefit when compared with either the static or zero baselines. This demonstrates the importance of constructing robust baselines rather than examining results out of context in benchmarking BP estimation.

## IV. DISCUSSION AND LIMITATIONS

Recent literature suggests how hypertension management could be revolutionized by ABPM and cuffless BP solutions—providing an alternative to in-office BP measurement that shortens time to diagnosis and elucidates missed white-coat or masked hypertension phenotypes [6]–[10]. However, these alternatives have not seen much uptake compared with in-office, cuff-based BP measurement. ABPM devices are notoriously uncomfortable, and many cuffless BP solutions have yet to be validated or fully assessed [3], [11]–[13].

Additionally, many studies of cuffless BP devices only compare device performance to an oscillometric reference device, and very few use auscultation as a reference [16]. As shown in our study, the performance of a wearable BP measurement device can be very different when compared to an oscillometric versus auscultatory reference. Thus, evaluating against an oscillometric reference alone may be misleading. Our single-timepoint return visit measurement results, which uniformly improve across feature groups and subpopulations in auscultatory rather than oscillometric reference conditions, suggest a strong decrease in noise with auscultation.

Our dataset is also more diverse and heterogeneous than many collected for cuffless BP estimation. We found the inclusion of older and hypertensive subpopulations critical to estimating real-world performance, as BP prediction performance was much stronger in young and normotensive subpopulations, likely due to comparatively smaller BP variability. Historically, cuffless BP devices have mainly evaluated performance based on small populations of young and healthy participants. Some recent cuffless BP studies have sought to address these performance discrepancies by recruiting larger and more diverse cohorts [14], [15]. It is perhaps of most importance that future studies include hypertensives and those with a large degree of BP variability. These population characteristics, together with ambulatory wear, render a more realistic BP measurement environment than recruiting a population that is merely diverse in age or gender.

Although this study demonstrates an evaluation of cuffless BP in the setting of its intended use and includes a diverse cohort of participants, there are known limitations. This study focuses on the relative differences among sensing techniques in predicting

BP changes, but the optical sensing device used does not necessarily represent the state of the art in PPG sensing. Despite this limitation, we expect the effects surrounding conditions of wear and study population to eclipse changes in sensing hardware. To showcase these conditions of wear and study design, we focused on building interpretable models with known features documented to be useful in BP prediction. Beyond the curated features included in our models, one could imagine developing a deep neural network model to learn waveform features without heuristics. This kind of model could yield predictive performance gains, but would sacrifice interpretability.

It is also important to note that our protocol, participant stratification, and analyses were not designed or evaluated to the FDA benchmark standard for evaluation of cuffless BP devices; thus, clinical utility cannot be determined from these data [27]. However, we still observe a stronger BP prediction performance in controlled as opposed to ambulatory settings. This is true even when using the same cuffless BP measurement techniques and identical devices, demonstrating the importance of evaluating BP measurement devices in the environment of their intended use.

To our knowledge, we present the largest dataset of wearable tonometry data. In this dataset, we observed that tonometry was the highest performing signal when it came to BP prediction—it came closest to approaching ANSI/AAMI validation standards and produced the most feature-rich waveform when compared with complementary PPG or ECG. While tonometry can provide a reduction in predicted BP bias and standard deviation, it is also less often available at a robust enough signal quality for feature extraction. In comparison with optical and ECG sensing techniques, tonometry is overall less comfortable and demands stricter user compliance in correct applanation technique. Innovations lowering the barriers for user compliance, such as measurement of tonometric data from a piezoresistive sensor patch, could increase availability of the tonometric signal.

## V. Conclusion

We have evaluated ABPM experimental protocols in settings of intended use and across a wide range of participants diverse in age, gender, BMI, and hypertensive status. Although wearable, cuffless BP devices are designed for continuous use in daily life, they are frequently only tested or validated in small, relatively healthy populations and against known measurement standards at rest. The importance of appropriate study design for the evaluation of wearable BP measurement was illustrated by: (1) the comparatively better performance of younger and normotensive subpopulations, (2) the strong performance of challenging benchmarks, including a static baseline regressed on initial in-lab BP, and (3) the difference in performance observed when approximating the intended use environment via 24-hour ambulatory wear.

The Aurora-BP study dataset represents the largest known collection of simultaneous tonometry, PPG, ECG, accelerometry, and BP measurements. It also provides the first example of semi-continuous tonometry over a 24-hour period. We observe that tonometry, when available, provides a more feature-rich waveform than PPG or ECG. Tonometry excels in the noise-reduced setting of sigle-timepoint BP measurement with auscultatory references and exhibits a low 24-hour average standard deviation in ambulatory settings. Differences in BP prediction performance between auscultatory and oscillometric references indicate the importance of robust BP references. This publicly available dataset will provide value to researchers aiming to gain further insights into cuffless BP predictive performance as well as patterns measured in both controlled and ambulatory settings.

## References


[1] National Health and Nutrition Examination Survey (NHANES) Data, Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS), 2022, [Online]. Available: https://www.cdc.gov/nchs/nhanes/

[2] F. Yang *et al.*, "Prevalence, awareness, treatment, and control of hypertension in the older population: Results from the multiple national studies on ageing," *J. Amer. Soc. Hypertension*, vol. 10, no. 2, pp. 140–148, 2016.

[3] B. B. Green *et al.*, "Blood pressure checks and diagnosing hypertension (BP-CHECK): Design and methods of a randomized controlled diagnostic study comparing clinic, home, kiosk, and 24-hour ambulatory BP monitoring," *Contemporary Clin. Trials*, vol. 79, pp. 1–13, 2019.

[4] W. B. White and V. Gulati, "Managing hypertension with ambulatory blood pressure monitoring," *Curr. Cardiol. Rep.*, vol. 17, no. 2, pp. 1–9, 2015.

[5] A. J. Viera, A. L. Hinderliter, A. V. Kshirsagar, J. Fine, and R. Dominik, "Reproducibility of masked hypertension in adults with untreated borderline office blood pressure: Comparison of ambulatory and home monitoring," *Amer. J. Hypertension*, vol. 23, no. 11, pp. 1190–1197, 2010.

[6] P. Muntner *et al.*, "Measurement of blood pressure in humans: A scientific statement from the american heart association," *Hypertension*, vol. 73, no. 5, pp. e35–e66, 2019.

[7] D. Shimbo, M. Abdalla, L. Falzon, R. R. Townsend, and P. Muntner, "Role of ambulatory and home blood pressure monitoring in clinical practice: A narrative review," *Ann. Intern. Med.*, vol. 163, no. 9, pp. 691–700, 2015.

[8] T. Fujiwara, Y. Yano, S. Hoshide, H. Kanegae, and K. Kario, "Association of cardiovascular outcomes with masked hypertension defined by home blood pressure monitoring in a Japanese general practice population," *JAMA Cardiol.*, vol. 3, no. 7, pp. 583–590, 2018.

[9] K. Asayama *et al.*, "Prediction of stroke by home "morning" versus "evening" blood pressure values: The ohasama study," *Hypertension*, vol. 48, no. 4, pp. 737–743, 2006.

[10] N. Kakaletsis *et al.*, "Prognostic value of 24-h ABPM in acute ischemic stroke for short-, medium-, and long-term outcome: A systematic review and meta-analysis," *Int. J. Stroke*, vol. 10, no. 7, pp. 1000–1007, 2015.

[11] D. Shimbo *et al.*, "The use of ambulatory blood pressure monitoring among medicare beneficiaries in 2007–2010," *J. Amer. Soc. Hypertension*, vol. 8, no. 12, pp. 891–897, 2014.

[12] S. T. Kent *et al.*, "Rates, amounts, and determinants of ambulatory blood pressure monitoring claim reimbursements among medicare beneficiaries," *J. Amer. Soc. Hypertension*, vol. 8, no. 12, pp. 898–908, 2014.

[13] I. M. Kronish *et al.*, "Barriers to conducting ambulatory and home blood pressure monitoring during hypertension screening in the United States," *J. Amer. Soc. Hypertension*, vol. 11, no. 9, pp. 573–580, 2017.

[14] A. Vybornova, E. Polychronopoulou, A. Wurzner-Ghajarzadeh, S. Fallet, J. Sola, and G. Wuerzner, "Blood pressure from the optical Aktiia Bracelet: A 1-month validation study using an extended ISO81060-2 protocol adapted for a cuffless wrist device," *Blood Press. Monit.*, vo. 26, no. 4, p. 305, 2021.

[15] D. Nachman *et al.*, "Comparing blood pressure measurements between a photoplethysmography-based and a standard cuff-based manometry device," *Sci. Rep.*, vol. 10, no. 1, pp. 1–9, 2020.

[16] K. Kario, D. Shimbo, N. Tomitani, H. Kanegae, J. E. Schwartz, and B. Williams, "The first study comparing a wearable watch-type blood pressure monitor with a conventional ambulatory blood pressure monitor on in-office and out-of-office settings," *J. Clin. Hypertension*, vol. 22, no. 2, pp. 135–141, 2020.

[17] F. Heydari, M. P. Ebrahim, J.-M. Redoute, K. Joe, K. Walker, and M. R. Yuce, "A chest-based continuous cuffless blood pressure method: Estimation and evaluation using multiple body sensors," *Inf. Fusion*, vol. 54, pp. 119–127, 2020.

[18] J. R. Sempionatto *et al.*, "An epidermal patch for the simultaneous monitoring of haemodynamic and metabolic biomarkers," *Nature Biomed. Eng.*, vol. 5, no. 7, pp. 737–748, 2021.

[19] N. Luo *et al.*, "Flexible piezoresistive sensor patch enabling ultralow power cuffless blood pressure measurement," *Adv. Funct. Mater.*, vol. 26, no. 8, pp. 1178–1187, 2016.

[20] E. Dewhirst *et al.*, "Accuracy of the cnap monitor, a noninvasive continuous blood pressure device, in providing beat-to-beat blood pressure readings in the prone position," *J. Clin. Anesth.*, vol. 25, no. 4, pp. 309–313, 2013.

[21] B. Silke and D. McAuley, "Accuracy and precision of blood pressure determination with the finapres: An overview using re-sampling statistics," *J. Hum. Hypertension*, vol. 12, no. 6, pp. 403–409, 1998.

[22] D. Stokes, T. Clutton-Brock, C. Patil, J. Thompson, and P. Hutton, "Comparison of invasive and non-invasive measurement of continuous arterial pressure using the finapres," *Brit. J. Anaesth.*, vol. 67, no. 1, pp. 26–35, 1991.

[23] J. S. Kim, K. K. Kim, H. J. Baek, and K. S. Park, "Effect of confounding factors on blood pressure estimation using pulse arrival time," *Physiol. Meas.*, vol. 29, no. 5, pp. 615–624, 2008.

[24] S. L.-O. Martin *et al.*, "Weighing scale-based pulse transit time is a superior marker of blood pressure than conventional pulse arrival time," *Sci. Rep.*, vol. 6, no. 1, pp. 1–8, 2016.

[25] A. Chandrasekhar, C.-S. Kim, M. Naji, K. Natarajan, J.-O. Hahn, and R. Mukkamala, "Smartphone-based blood pressure monitoring via the oscillometric finger-pressing method," *Sci. Transl. Med.*, vol. 10, no. 431, 2018, Art. no. eaap8674.

[26] C.-S. Kim, A. M. Carek, O. T. Inan, R. Mukkamala, and J.-O. Hahn, "Ballistocardiogram-based approach to cuffless blood pressure monitoring: Proof of concept and potential challenges," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 11, pp. 2384–2391, Nov. 2018.

[27] *Non-Invasive Sphygmomanometers – Part 2: Clinical Investigation of Automated Measurement Type*, Standard ISO 81060-2, Association for the Advancement of Medical Instrumentation, Arlington, VA, Jun. 2013.

[28] R. Mukkamala *et al.*, "Toward ubiquitous blood pressure monitoring via pulse transit time: Theory and practice," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 8, pp. 1879–1901, Aug. 2015.

[29] C. Poon and Y. Zhang, "Cuff-less and noninvasive measurements of arterial blood pressure by pulse transit time," in *Proc. IEEE Eng. Med. Biol. 27th Annu. Conf.*, 2006, pp. 5877–5880.

[30] C. Yang and N. Tavassolian, "Pulse transit time measurement using seismocardiogram, photoplethysmogram, and acoustic recordings: Evaluation and comparison," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 3, pp. 733–740, May 2018.

[31] X. Ding, B. P. Yan, Y.-T. Zhang, J. Liu, N. Zhao, and H. K. Tsang, "Pulse transit time based continuous cuffless blood pressure estimation: A new extension and a comprehensive evaluation," *Sci. Rep.*, vol. 7, no. 1, pp. 1–11, 2017.

[32] C. Vlachopoulos, M. O'Rourke, and W. W. Nichols, *McDonald's Blood Flow in Arteries: Theoretical, Experimental and Clinical Principles*. Boca Raton, FL, USA: CRC Press, 2011.

[33] R. Mukkamala and J.-O. Hahn, "Toward ubiquitous blood pressure monitoring via pulse transit time: Predictions on maximum calibration period and acceptable error limits," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 6, pp. 1410–1420, Jun. 2018.

[34] P. K. Whelton *et al.*, "2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: A report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines," *J. Amer. College Cardiol.*, vol. 71, no. 19, pp. e127–e248, 2018.

[35] C.-H. Chen *et al.*, "Estimation of central aortic pressure waveform by mathematical transformation of radial tonometry pressure: Validation of generalized transfer function," *Circulation*, vol. 95, no. 7, pp. 1827–1836, 1997.

[36] B. Fetics, E. Nevo, C.-H. Chen, and D. A. Kass, "Parametric model derivation of transfer function for noninvasive estimation of aortic pressure by radial tonometry," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 6, pp. 698–706, Jun. 1999.

[37] X. Peng, M. G. Schultz, W. P. Abhayaratna, M. Stowasser, and J. E. Sharman, "Comparison of central blood pressure estimated by a cuff-based device with radial tonometry," *Amer. J. Hypertension*, vol. 29, no. 10, pp. 1173–1178, 2016.

[38] P. Digiglio, R. Li, W. Wang, and T. Pan, "Microflotronic arterial tonometry for continuous wearable non-invasive hemodynamic monitoring," *Ann. Biomed. Eng.*, vol. 42, no. 11, pp. 2278–2288, 2014.

[39] L. Garcia-Ortiz *et al.*, "Noninvasive validation of central and peripheral augmentation index estimated by a novel wrist-worn tonometer," *J. Hypertension*, vol. 36, no. 11, pp. 2204–2214, 2018.

[40] E. O'Brien, G. Parati, and G. Stergiou, "Ambulatory blood pressure measurement: What is the international consensus?," *Hypertension*, vol. 62, no. 6, pp. 988–994, 2013.

[41] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Trans. Biomed. Eng.*, no. 3, pp. 230–236, Mar. 1985.

[42] O. Vardoulis *et al.*, "In vivo evaluation of a novel, wrist-mounted arterial pressure sensing device versus the traditional hand-held tonometer," *Med. Eng. Phys.*, vol. 38, no. 10, pp. 1063–1069, 2016.

[43] J. P. Murgo, N. Westerhof, J. P. Giolma, and S. A. Altobelli, "Aortic input impedance in normal man: Relationship to pressure wave forms," *Circulation*, vol. 62, no. 1, pp. 105–116, 1980.

[44] X.-R. Ding, Y.-T. Zhang, J. Liu, W.-X. Dai, and H. K. Tsang, "Continuous cuffless blood pressure estimation using pulse transit time and photoplethysmogram intensity ratio," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 5, pp. 964–972, May 2016.

[45] W. Chen, T. Kobayashi, S. Ichikawa, Y. Takeuchi, and T. Togawa, "Continuous estimation of systolic blood pressure using the pulse arrival time and intermittent calibration," *Med. Biol. Eng. Comput.*, vol. 38, no. 5, pp. 569–574, 2000.

[46] Y.-L. Zheng, B. P. Yan, Y.-T. Zhang, and C. C. Poon, "An armband wearable device for overnight and cuff-less blood pressure measurement," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 7, pp. 2179–2186, Jul. 2014.

[47] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006, ch. 4, pp. 179–224.

[48] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*. Boca Raton, FL, USA: CRC Press, 1994.

[49] G. D. Ruxton, "The unequal variance *t*-test is an underused alternative to Student's *t*-test and the Mann-whitney U test," *Behav. Ecol.*, vol. 17, no. 4, pp. 688–690, 2006.