Ametris
A SIGNANT HEALTH COMPANY

TECHNICAL WHITE PAPER

Device Comparability and Data Compatibility Across ActiGraph Generations

Technical White Paper

Device Comparability and Data Compatibility Across ActiGraph Generations

by Samuel R. LaMunion, PhD

As the digital health landscape continues to evolve, researchers are increasingly looking to maximize the power of their data by pooling datasets across multi-year longitudinal studies, or by combining independent datasets across different cohorts within a specific disease or therapeutic area. **However, a critical question keeps coming up:**

> *How comparable are physical behavior measurements across different device generations/models?*

Whether you are harmonizing data over a decade-long study or attempting to aggregate data from different clinical trials evaluating a specific disease area, minor variations in device performance or characteristics can subtly alter how we interpret physical activity, sedentary time, and sleep at both the individual and group levels. To address this, our team set out to establish a standardized, reproducible verification protocol to test device comparability directly.

This original research, now published in Medicine & Science in Sports & Exercise (MSSE), represents a multi-year collaborative effort between the National Institutes of Health (NIH) and Ametris (formerly ActiGraph).

Why Sensor Verification Matters:

The V3+ Framework

The intent of this study was rooted in verification, a core element of the Digital Medicine Society's (DiMe) V3+ framework. While sensor verification is highly technical and takes place **"in the weeds,"** it is an essential first step. Before we can validate a digital clinical measure or use it in a trial, we must have absolute confidence that the underlying sensor hardware performs consistently.

Historically, evaluating sensor performance was limited by the lack of standardization in testing environments. To overcome this, our protocol relied heavily on a customized, National Institute of Standards and Technology (NIST)-traceable orbital shaker. Utilizing a mechanical shaker introduces several critical advantages for hardware verification:

- **Highly Reproducible Accelerations**
  Unlike human movement, which varies naturally from second to second and person to person, a mechanical shaker delivers a fixed, consistent frequency profile that can be identically replicated across multiple testing runs.

- **Simultaneous Multi-Device Testing**
  In a human trial, you can only realistically place a few devices on a single participant at once to capture true human behavior. The shaker allows us to test a large batch of unique devices simultaneously under identical conditions, giving us the statistical power to capture actual within and between-generation hardware variability.

- **Mitigation of Behavioral Noise**
  By using custom trays and tight foam inserts to secure the devices, we isolated the precise physical acceleration of the shaker. This completely removed the confounding artifact noise and unpredictable movement variations inherent to human wearers, leaving us with pure, unadulterated hardware performance data.

With those characteristics in mind, we designed our study to establish a standardized, reproducible protocol to evaluate accelerometer-based wearable sensors. Our first demonstration of this protocol was to evaluate five generations of devices ranging from the legacy GT3X+ (used in the 2011-2014 cycles of the U.S. population-representative National Health And Nutrition Examination Survey [NHANES]) up to the CenterPoint InsightWatch v2.0. Using a custom, National Institute of Standards and Technology (NIST)-traceable orbital shaker, we evaluated 15 unique devices per device generation with the lone exception being an n = 11 for the CPIWv2 due to availability at the time of testing. Devices were evaluated against a gold-standard, NIST-certified criterion accelerometer across 11 different shaker frequencies equivalent to a physiologically relevant dynamic acceleration range of 0 to 3.7 g's. For reference, in free-living students of both youth and adults, average sustained accelerations for vigorous activities were between 2-3 g's.

Key Findings:

Static Offset, Dynamic Response, and Signal Drift

We focused on characterizing three distinct components of sensor behavior:

- **Static Offset (i.e., bias from 1 g vector magnitude during periods of non-movement)**
  This measures a device's deviation from a 1 g vector magnitude at rest. Older device generations (GT3X+) exhibited the highest within-generation variability, with some units exceeding the target of ±50 milli-g's error threshold. Newer device generations were substantially tighter and more consistent. We also noted an axis-dependence (XY vs. XZ orientations) due to different manufacturing tolerances for the Z-axis, whereas user-selected sampling frequency (30 vs 80 Hz or 32 vs 64 Hz depending on the device generation) had no observable or statistical impact.

- **Dynamic Response (i.e., In-Motion Performance)**
  Device acceleration response was largely statistically equivalent across all generations under both absolute (±50 milli-g's) and relative (±5%) criteria. Minor deviations occurred only at the highest acceleration forces (~3 g's) for the wGT3X-BT and GT9X. Because this occurs well within the vigorous physical activity range, the real-world impact of this variance is likely inconsequential.

- **Signal Drift (i.e., Stability Over Time)**
  Over the four-hour testing cycles, baseline signal drift was negligible across all generations, meaning the sensor baselines remained stable.

The "So What?"

Real-World Impact on Free-Living Data

To understand the practical implications of these hardware differences, we ran a simulation using free-living data from NHANES participants at the 25th, 50th, and 75th activity percentiles.

- **The Worse Case Scenario**
  We artificially applied the worst empirically derived static offset to this real-world data to see how it would affect downstream metrics if left entirely uncorrected.

The results of this experiment were telling. In the raw acceleration data, the uncorrected static offset caused massive errors resulting in mean absolute percent errors (MAPE) in the hundreds of percent for total activity volume (Euclidean Norm Minus One [ENMO]), minutes of moderate-to-vigorous physical activity (MVPA), and inactivity. Interestingly, sleep estimates remained largely unaffected.

A Solution

Autocalibration is Highly Effective

An encouraging takeaway from our work is that common data processing techniques can elegantly solve these hardware discrepancies, to an extent.

When we applied autocalibration (van Hees 2014), a post-processing step which uses stationary intervals during free-living wear to calculate and adjust for individual axis gains and offsets, the between-device variability collapsed to effectively zero.

- **A Practical Consideration for Researchers**
  While autocalibration works fairly seamlessly during multi-day, free-living studies using a wrist-worn device where stationary periods are abundant, it can fail in shorter laboratory validation studies (2–6 hours) because the device doesn't collect enough multi-axis stationary data to adequately populate a 3-dimensional sphere of data points across the triaxial device.

For short lab studies, we recommend a hybrid calibration approach: pause or start the study by deliberately placing the device at known, static angles relative to gravity, similar to how factory calibration coefficients may be obtained. This creates the necessary baseline data to support accurate autocalibration within a collection period.

Final

Takeaways

1

A Reproducible Framework

We have successfully introduced an objective, open framework for researchers to compare accelerometer performance across different generations and manufacturers.

2

Harmonization is Achievable

While hardware differences across device generations exist, they do not inherently break data compatibility.

3

Calibration is Critical

Software processing solutions like autocalibration are not just minor technical details, they are scientifically necessary to ensure not only data harmonization but also accurate clinical interpretations.

This research was originally presented at **ADDS 2026 in Atlanta, GA**.

You can view the full recorded presentation here or read the complete peer-reviewed study in MSSE.

Samuel R. LaMunion, PhD

Dr. LaMunion is a quantitative physiologist and research scientist specializing in digital health technologies, physical behavior assessment, and wearable sensor calibration and validation methodologies. This research was completed during his time as a Postdoctoral Fellow within the Intramural Research Program at the National Institutes of Health (NIH), National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK).

sam@lrs-c.com | linkedin.com/in/samuel-lamunion

Disclosures

1. The views presented in this article and the accompanying presentation are entirely my own and do not constitute an endorsement by the National Institutes of Health or the U.S. government.
2. Samuel R. LaMunion has engaged in paid consulting work for Ametris, however, the entirety of this work pre-dates that affiliation.

Acknowledgements

This technical white paper is an operational summary of a multi-year, collaborative research initiative. I would like to acknowledge and thank my collaborators, the core study co-authors and investigators, for their invaluable contributions to the underlying research design and data collection and for providing the resources to make this project possible: Joe Nguyen (Ametris), Dr. Robert J. Brychta (National Institutes of Health [NIH], National Institute of Diabetes, Digestive, and Kidney Diseases [NIDDK]), Dr. Richard P. Troiano (United States Public Health Service, retired), Dr. Karl E. Friedl (United States Army Research Institute of Environmental Medicine [USARIEM]), and Dr. Kong Y. Chen (NIH/NIDDK).

References

1. **Primary Publication:** LaMunion, S. R., et al. (2025). Device Comparability and Data Compatibility Across ActiGraph Generations. Medicine & Science in Sports & Exercise, Jun 1;57(6):1161-1173. https://doi.org/10.1249/MSS.0000000000003652.
2. **Methodological Reference:** van Hees, V. T., et al. (2014). Autocalibration of accelerometer data for free-living physical activity assessment. Journal of Applied Physiology, 117(7), 738-744. https://doi.org/10.1152/japplphysiol.00421.2014.
3. **Framework Reference:**
   - Goldsack, J. C., et al. (2020). Verification, analytical validation, and clinical validation (V3): the foundation of determining fit-for-purpose for Biometric Monitoring Technologies (BioMeTs). NPJ digital medicine, 3, 55. https://doi.org/10.1038/s41746-020-0260-4.
   - Bakker, J. P., et al. (2025). V3+ extends the V3 framework to ensure user-centricity and scalability of sensor-based digital health technologies. NPJ digital medicine, 8(1), 51. https://doi.org/10.1038/s41746-024-01322-2

Ametris
A SIGNANT HEALTH COMPANY

Founded in 2004, Ametris (formerly ActiGraph) is a global digital health solutions provider transforming real-world patient data into validated clinical evidence. Our end-to-end platform combines advanced wearables, regulatory-aligned analytics, and expert scientific support across every phase of a clinical trial—from protocol design to regulatory submission. We serve leading pharmaceutical, biotech, and public health research organizations worldwide, backed by two decades of experience and a collaborative, science-first approach.

In 2026, Ametris joined Signant Health, uniting wearable-derived digital outcome measures with industry-leading eCOA solutions to give sponsors a single partner for integrated evidence generation. Together, we combine continuous sensor data with patient-reported outcomes for a more complete picture of patient health.