A Multi-Source Validation Framework for Smartphone-Based Walk Tests

Abreu, M[1]; Clancy, C[1]; Lutz, J[1,2,3]
[1] Beats Medical, Dublin, Ireland
[2] Boston University, Chobanian & Avedisian School of Medicine, Boston, USA
[3] Luja Consulting LLC, Pittsfield, USA

Methodological Issue Being Addressed

Evaluating smartphone-based walk tests as digital mobility endpoints requires evidence from multiple complementary sources rather than a single validation study.

A fit-for-purpose assessment must integrate complementary evidence spanning technical performance, benchmarking, cross-platform consistency, and external validation. Here we address the methodological problem of how to organize these distinct evidence sources into a coherent framework for judging readiness of smartphone-based 2-Minute Walk Test (2MWT) and 6-Minute Walk Test (6MWT) assessments.

Introduction

The 2MWT and 6MWT are widely used assessments of functional mobility across neurological and chronic disease populations. These assessments are traditionally conducted in person using standardized walking courses and trained personnel, creating logistical and operational burdens for large-scale or remote studies. Smartphone-based implementations offer a scalable, low-burden approach to remote mobility assessment using participants' own devices and may facilitate more frequent and geographically distributed mobility assessments. However, establishing confidence in these measures requires evidence from multiple complementary evaluation approaches.

Methods

1 External Benchmarking

Exploratory comparison against Mobilise-D gait-speed references (1)

Descriptive comparison of algorithm-derived gait speed from one participant per cohort against published cohort-level reference values.

Analysis Directional alignment

2 Cross-Platform Harmonization

Simultaneous iOS and Android recordings

Healthy adults (n=2, 5 recordings)
Adults with Parkinson's disease (n=3, 9 recordings).

Analysis Bland-Altman analyses and percentage differences

3 Analytical Validation

Smartphone-derived 6MWT vs video reference

Healthy adults (n=6, 8 recordings).

Analysis MAPE with 95% CI

4 Independent Validation

Smartphone-derived 2MWT vs manual reference (2)

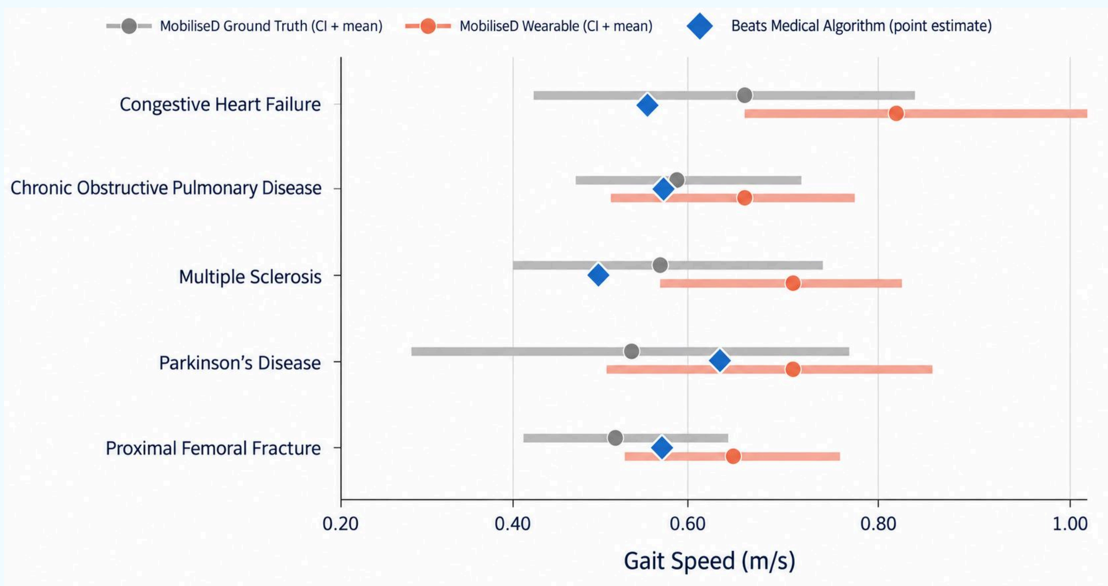
Independently collected 2MWT data from healthy adults (n=37)

Analysis Bland-Altman and correlation analyses

Results

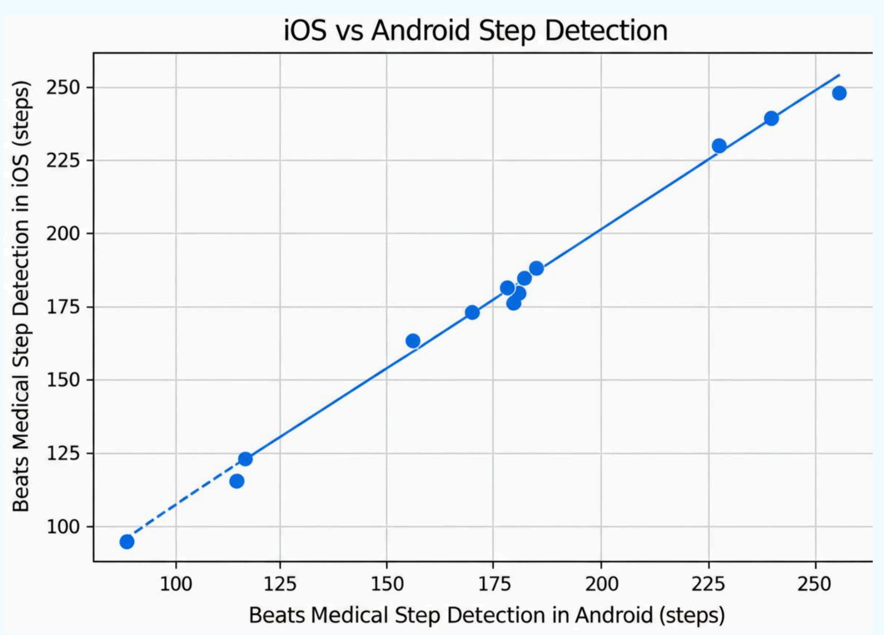
1

External Benchmarking: Algorithm-derived gait-speed estimates were directionally aligned with Mobilise-D cohort-level reference values, remaining numerically close to the published mean.


MobiliseD Ground Truth (CI + mean)
MobiliseD Wearable (CI + mean)
Beats Medical Algorithm (point estimate)
Congestive Heart Failure
Chronic Obstructive Pulmonary Disease
Multiple Sclerosis
Parkinson's Disease
Proximal Femoral Fracture
0.20
0.40
0.60
0.80
1.00
Gait Speed (m/s)

2

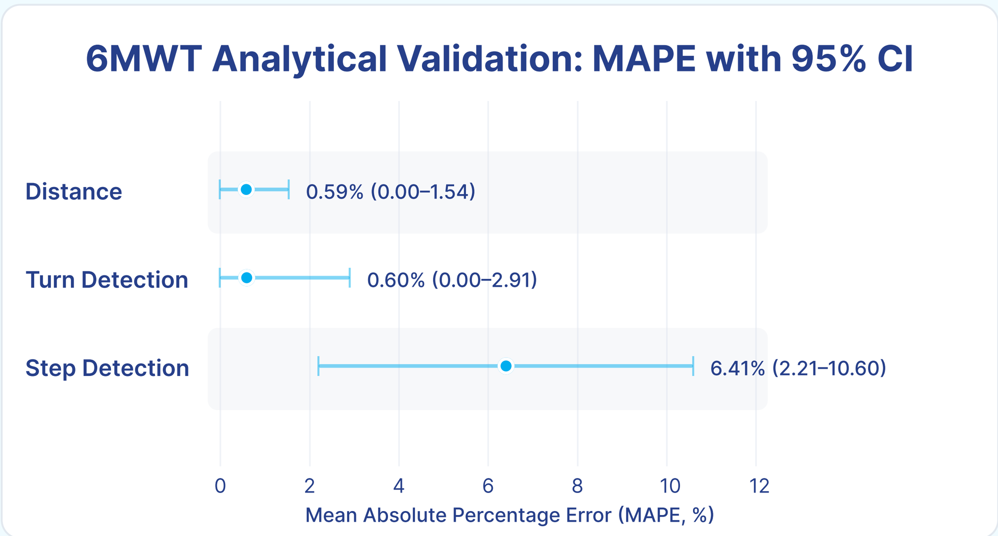
Cross-Platform Harmonization: Consistent step detection was observed across iOS and Android recordings, with a Bland–Altman bias of 0.21 steps and all observations within the limits of agreement (−6.13 to 6.56 steps). The mean percentage difference was 1.56% (95% CI: 0.36–2.94%), with a strong correlation between platforms (Spearman ρ=0.99, p<0.001).


iOS vs Android Step Detection
Beats Medical Step Detection in iOS (steps)
Beats Medical Step Detection in Android (steps)

3

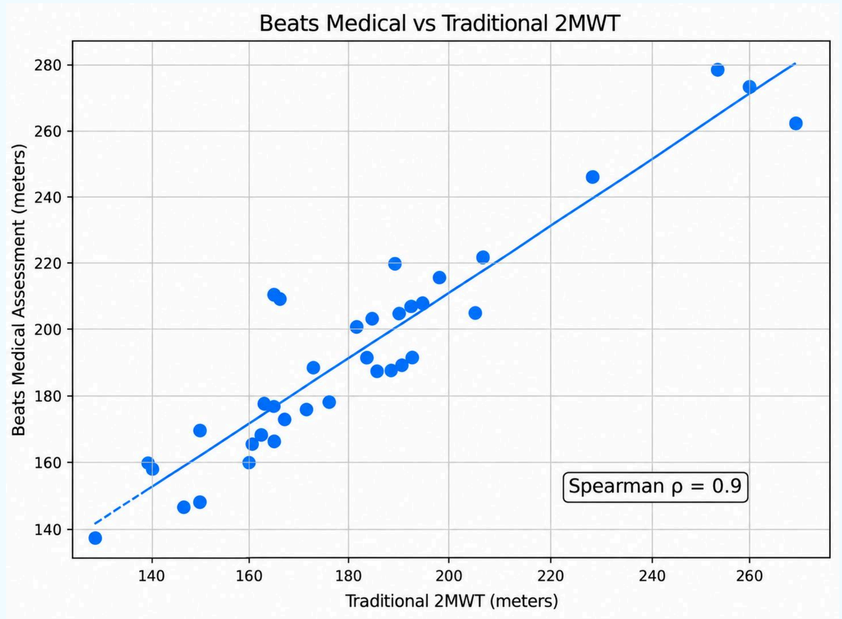
Analytical validation:

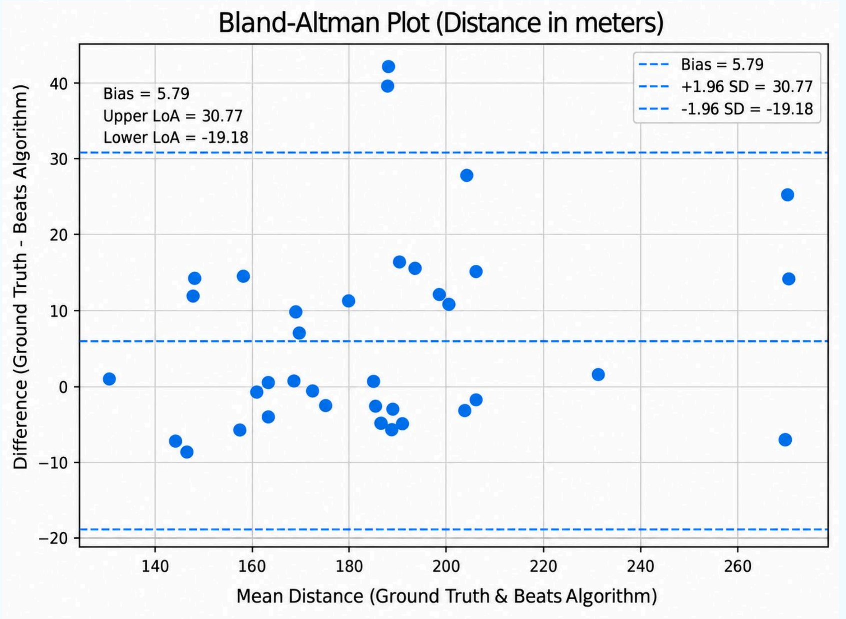
6MWT outcomes showed low error relative to video reference.
MAPE Distance: 0.59% (95% CI: 0.00-1.54)
MAPE Turn Detection: 0.60% (95% CI: 0.00-2.91)
MAPE Step Detection: 6.41% (95% CI: 2.21-10.60).


6MWT Analytical Validation: MAPE with 95% CI
Distance
0.59% (0.00–1.54)
Turn Detection
0.60% (0.00–2.91)
Step Detection
6.41% (2.21–10.60)
Mean Absolute Percentage Error (MAPE, %)

4

Independent Validation Smartphone-derived 2MWT: Distance bias: 5.79 m (LoA: -19.18 to 30.77 m, 94.6% of samples within the LoA). Turn detection bias: 0.41 turns (LoA: −1.16 to 1.97; 91.9% within LoA). Strong correlations were observed for distance (ρ=0.88, p<0.001) and turn counts (ρ=0.90, p<0.001).


Beats Medical vs Traditional 2MWT
Beats Medical Assessment (meters)
Traditional 2MWT (meters)
Spearman ρ = 0.9


Bland-Altman Plot (Distance in meters)
Bias = 5.79
Upper LoA = 30.77
Lower LoA = -19.18
Bias = 5.79
+1.96 SD = 30.77
-1.96 SD = -19.18
Difference (Ground Truth - Beats Algorithm)
Mean Distance (Ground Truth & Beats Algorithm)

Conclusions

• A multi-source validation framework can provide a structured approach for evaluating smartphone-based mobility endpoints.
• Convergent evidence from benchmarking, cross-platform harmonization, analytical validation, and independent validation supports the fit-for-purpose use of smartphone-derived 2MWT and 6MWT.
• This framework guides evidence generation and readiness assessments for digital mobility measures intended for clinical research and future clinical validation.

Limitations

• Evidence sources included modest and heterogeneous sample sizes.
• External benchmarking was based on published cohort-level reference values, and cross-platform harmonization was evaluated in a limited number of recordings.

References:

1 Kirk C, et al. Mobilise-D insights to estimate real-world walking speed in multiple conditions with a wearable device. Sci Rep. 2024;14:1754. doi:10.1038/s41598-024-51766-5.

2 Trinity College Dublin, Department of Physiotherapy. An Investigation into the Validity of a Smartphone-Based Two-Minute Walk Test Compared to the Traditional Method in a Healthy Population. Unpublished Final Year Project; 2023.

Acknowledgements/Disclosures: This work was funded by Beats Medical. One or more authors report potential conflicts which are described in the program.

BEATS MEDICAL