OneStep

WHITE PAPER · PASSIVE GAIT MONITORING

How Humans Walk Differently When Conscious vs Unconscious of Their Movement

Comparing in-app and background smartphone gait measurement, and what background data adds to clinical monitoring

ONESTEP RESEARCH | AUGUST 2023 | ONESTEP.CO/RESOURCES

RE How Humans Walk Differently When Conscious vs Unconscious of Their Movement

The walk a patient shows you is not the walk they take

Almost every gait assessment ever made was taken by someone who knew they were being measured. OneStep's background monitoring records the other walk: the one patients take through their day, with the phone in their pocket, without pressing a button.

330	11	2×	77
Patients with both in-app and background walks	Gait parameters compared	More monitoring coverage with background data	Joint-replacement recoveries tracked both ways

This paper asks a simple question with large consequences for care: does the natural, unconscious walk match the one patients record on purpose? Mostly it does. After the right aggregation, in-app and background measurements agree well for most parameters. But the background walk is consistently a little less healthy, a few patients walk very differently when nobody is watching, and recovery after hip and knee replacement takes visibly longer in everyday walking than conscious recordings suggest.

What this means for a care team

See the real recovery

Prompted walks say a patient is back to baseline weeks before their everyday walking agrees. Background data lets you know when a patient is truly back to normal.

Catch what in-app misses

Background monitoring found 156 declines in 59 patients at times when there was no in-app data to analyze. For 39 of them, in-app recordings never flagged anything.

Double the coverage, no extra effort

Two-week periods with enough data for a change-of-status check nearly doubled, from 1,102 to 2,184, without asking patients to remember to record.

HOW TO ACT ON IT

The paper proposes a simple rule: a change detected by one source alone is a **yellow alert**; a change confirmed by both in-app and background data, or by other corroborating evidence, is a **red alert**. It is the difference between knowing a patient's best walk and knowing their real one.

© 2026 OneStep

The full white paper follows.

WHITE PAPER · AUGUST 2023

How Humans Walk Differently When Conscious vs Unconscious of Their Movement

Agreement between in-app and background smartphone gait measurement, and the clinical use of passively collected gait data

OneStep Research Team
OneStep · Correspondence: info@onestep.co

ABSTRACT

Background. Nearly all gait assessment is performed while the person is aware of being measured. Awareness may alter gait, so consciously recorded walking may not represent a person's walking through the day. OneStep's smartphone gait analysis can record walks passively ("background" walks) as well as on demand ("in-app" walks).

Objective. To determine how background gait data should be aggregated and interpreted, how closely it agrees with in-app measurement, and whether it can support the same clinical uses: tracking recovery after joint replacement and detecting meaningful change in gait.

Methods. Retrospective analysis of OneStep's gait database. The primary cohort comprised 330 patients with at least 10 in-app and 25 background recordings. Eleven gait parameters were studied. Six daily aggregation methods were compared by within-week variance. Weekly in-app and background values were compared with intraclass correlation coefficients ($ICC_{3,1}$) and Bland–Altman analysis. Recovery trajectories were compared in 45 total hip and 32 total knee replacement patients. Change-of-status detection from the two sources was compared with chi-squared tests.

Results. The best aggregation method differed by parameter. After aggregation, 6 of 11 parameters showed good agreement (ICC ≥ 0.75) and two more ICC ≥ 0.70; consistency and step length asymmetry agreed poorly (ICC < 0.5). Background values were shifted toward less healthy gait for all parameters. A small number of patients showed large in-app–background differences. Background walks tracked post-surgical recovery but returned to pre-surgery levels more slowly than in-app walks. Change-of-status detection from baseline was strongly associated between sources for improvement ($p = 6.5 \times 10^{-24}$) and decline ($p = 5.6 \times 10^{-7}$). Background data nearly doubled the periods eligible for change-of-status analysis (2,184 vs 1,102; +98%) and detected 156 declines in 59 patients when in-app data were insufficient.

Conclusions. Properly aggregated background gait data agree well with in-app data for most parameters while capturing the everyday walk that conscious recording misses. Background monitoring extends clinical visibility, reveals that real-world recovery is slower than prompted tests suggest, and supports a two-tier alerting approach.

KEY POINTS

Question	Does the gait people record on purpose match the gait they walk with through the day, and can passively collected gait data support clinical monitoring?
Findings	In 330 patients, in-app and background gait agreed well for most parameters after aggregation, but background gait was consistently slightly less healthy; recovery after hip and knee replacement took longer in background walks; background data doubled monitoring coverage and found declines invisible to in-app recording.
Meaning	Background gait monitoring complements conscious measurement: it shows when a patient is truly recovered and detects change without relying on patients to record themselves.

KEYWORDS gait analysis · passive monitoring · smartphone · remote therapeutic monitoring · joint replacement · change detection · Bland–Altman · ICC

© 2026 OneStep · www.onestep.co

1 Introduction

Gait analysis is a valuable tool for understanding health, fitness and mobility, and during physical therapy in particular.[1–3] Methods range from pressure-sensitive mats and treadmills to wearable devices,[4] but almost all share one feature: they measure gait as an active process, with the subject aware that their walking is being analyzed.

Conscious awareness of walking may itself affect gait, so a consciously recorded walk does not necessarily represent a person's gait through the day when they are not attending to how they walk.[5–7] Measuring gait in the natural environment is therefore an important resource for clinical outcomes.[8,9] Recent technology allows gait to be recorded at home,[10] including OneStep's smartphone gait analysis, which yields clinically validated parameters from a phone in the pocket.[11–13] Even so, when a patient presses a button to record a walk, they know their gait is being measured. Patients in physical therapy may make a special effort to walk well while recording themselves yet show poor mechanics when less focused.

OneStep's background gait analysis addresses this gap. With the application running in the background and the phone in the pocket, walks are recorded without the patient choosing to start a recording (background recording is enabled by the user when prompted and can be disabled). This captures truly unconscious, natural gait. Because such data are recorded across many conditions and contexts, they are noisier than in-app recordings made under ideal conditions. Here we present methods to reduce this noise, compare background with in-app gait, examine the people whose two gaits differ most, and test background data in two clinical uses for which in-app analysis is already applied: recovery after joint replacement and detection of meaningful change in gait over time.

2 Methods

2.1 Data source and cohort

All data came from OneStep's gait database of patients who used OneStep as part of rehabilitation. The primary analysis included patients with at least 10 in-app and 25 background walk measurements; 330 patients met this criterion.

2.2 In-app and background recordings

An in-app recording begins when the user opens the application, selects a short or long walk, presses start and walks with the phone in the pocket until a timer ends or the recording is stopped manually. A background recording occurs when a patient who has enabled background recording walks with the phone in the trouser pocket without initiating a recording: the application periodically checks whether the phone's movement is consistent with walking with the phone in a pocket and, if so, records and stores a 60-second measurement labeled as background. Both recording types are processed by the same pipeline: recordings are segmented into strides, pre-processed into the circularized repetitive inertial pattern of each stride, and passed to machine-learning models that extract gait parameters.

2.3 Gait parameters

Eleven parameters were studied: cadence (steps/min), velocity (m/s), double support (% of time with both feet on the ground), stance left and right (% of time the leg is in contact with the ground), stance asymmetry (left–right difference in stance), step length left and right (cm, heel to heel), step length asymmetry (left–right difference as % of stride length), stride length (cm) and consistency (similarity of successive steps on a 0–100 scale).

2.4 Data cleaning

Walk times were adjusted to the local time zone and a "day" was defined as 3:00 AM to 3:00 AM. Walks not fully analyzed by the algorithms (for example, phone placement other than a trouser pocket, too few steps, or movement not recognized as walking) were excluded. Weeks containing surgery or another acute medical event were excluded from weekly aggregation. Weekly values were generated only for weeks with background walks on at least 5 days and at least 3 in-app recordings.

2.5 Aggregation of background data

To obtain a representative daily value for each parameter, six aggregation methods were tested: mean, median, maximum, minimum, 25th percentile and 75th percentile. For each method and parameter, daily values were computed and the standard deviation of those values within each week was calculated. The method with the lowest average weekly standard deviation (averaged over values up to the 95th

© 2026 OneStep · www.onestep.co

percentile to limit the influence of outliers) was selected.

2.6 Statistical analysis

Weekly in-app values (mean of all in-app walks in the week) and weekly background values (mean of the aggregated daily values) were compared with $ICC_{3,1}$, treating the two recording modes as two raters of the same patient's gait, and with Bland–Altman plots. ICC was interpreted as poor (< 0.5), moderate (0.5–0.75), good (0.75–0.9) or excellent (> 0.9).[14] Associations between in-app and background change-of-status detections were tested with chi-squared tests on contingency tables. Analyses used the scipy and pingouin Python packages.

2.7 Patients with large in-app–background differences

For six parameters (double support, gait speed, stride length, left stance, stance asymmetry and step length asymmetry), patients were identified who had at least one week in which the difference between mean in-app and mean background values exceeded a parameter-specific cutoff (Table S1). Cases were then inspected manually.

2.8 Post-surgical recovery

Patients with both in-app and background measurements who underwent total hip replacement (THR; n = 45) or total knee replacement (TKR; n = 32) while using OneStep were analyzed. For each patient and each post-operative week, weekly averages were computed (after daily aggregation for background walks); population means and standard errors were then computed per week.

2.9 Change of gait status

Change of status between two time periods was assessed for five parameters: for each, the difference in means was tested with a one-directional t-test, the p-values were combined into a single value, and a change was declared when the combined $p < 0.05$ and at least one parameter changed by more than a threshold of clinical significance. Separate tests were run for improvement and for decline. Two designs were examined: short-term change between consecutive two-week periods, and change from baseline, comparing each two-week period with the patient's first two-week period (excluding periods containing surgery or another acute event).

3 Results

3.1 Reducing noise in background data

Background data had higher variance across walking bouts than in-app data, reflecting varied environments, footwear, surfaces and types of walking. The aggregation method that minimized weekly variance differed by parameter (Figure 1, Table 1; Figure S1): for example, the maximum for consistency, the 75th percentile for cadence, the 25th percentile for double support and stance, and the mean for velocity, stride and step lengths.

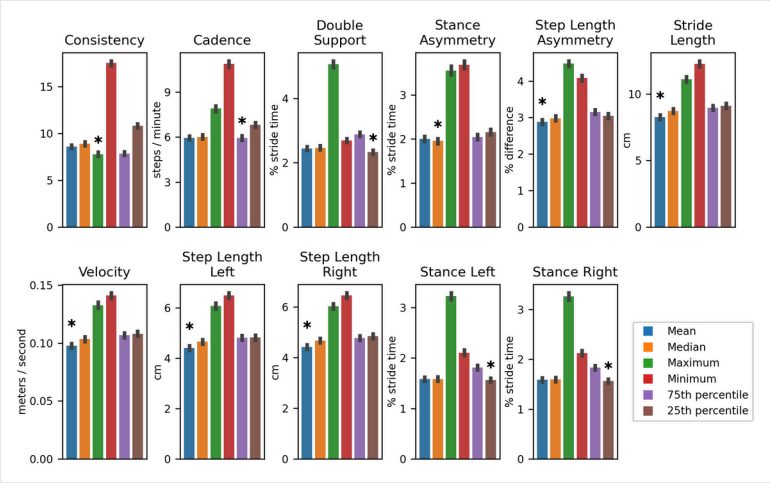


Figure 1. Weekly standard deviation of six daily aggregation methods, by parameter. For each parameter, a daily aggregate was computed by each method and the standard deviation of aggregates within each week was averaged. Stars mark the method with the lowest weekly standard deviation.

Table 1. Recommended daily aggregation method for background walks.

Parameter	Aggregation method	Parameter	Aggregation method
Consistency	Maximum	Velocity	Mean
Cadence	75th percentile	Step length left	Mean
Double support	25th percentile	Step length right	Mean
Stance asymmetry	Median	Stance left	25th percentile
Step length asymmetry	Mean	Stance right	25th percentile
Stride length	Mean		

3.2 Agreement between in-app and background values

After aggregation, weekly in-app and background values were compared within patient and week (Figure 2; Figure S2). Agreement varied by parameter (Table 2). Six of eleven parameters showed good reliability (ICC ≥ 0.75), two more reached ICC ≥ 0.70, while consistency and step length asymmetry showed poor reliability (ICC < 0.5).

Background values were systematically shifted toward the direction associated with less healthy gait: background cadence was lower than in-app cadence, and background double support was higher than in-app double support, with equivalent shifts in the other parameters. This confirms, in a larger dataset, the earlier finding of Rozanski and Putrino.[7]

Table 2. Agreement between weekly in-app and background values ($ICC_{3,1}$), from Figure 2 and Figure S2.

Parameter	ICC	Reliability	Parameter	ICC	Reliability
Double support	0.803	Good	Velocity	0.727	Moderate
Stance left	0.790	Good	Stance right	0.721	Moderate
Cadence	0.785	Good	Stance asymmetry	0.662	Moderate
Stride length	0.768	Good	Step length asymmetry	0.364	Poor
Step length left	0.758	Good	Consistency	0.321	Poor
Step length right	0.757	Good			

A Cadence **B Double support**

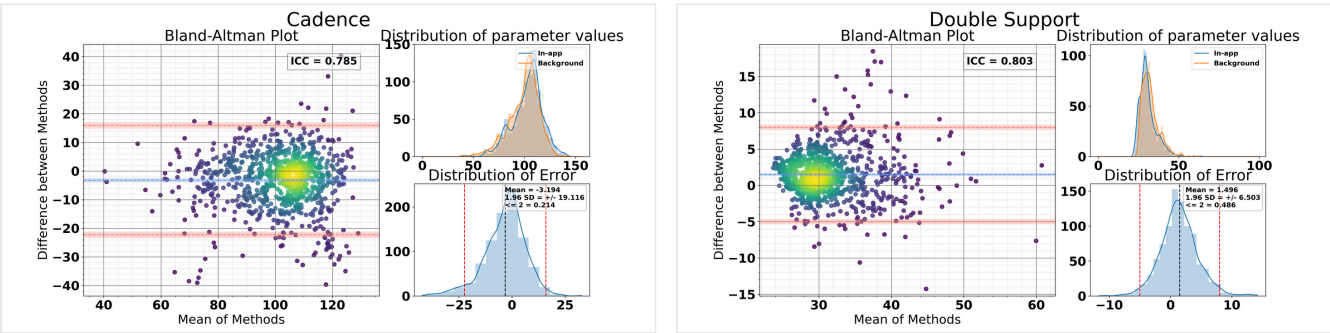


Figure 2. Comparison of weekly in-app and background values for cadence (A) and double support (B). In-app weekly values are the mean of all in-app walks in the week; background weekly values are the mean of daily aggregates computed with the method in Table 1. Left of each panel: Bland–Altman plot with $ICC_{3,1}$; the dotted blue line is the mean difference. Upper right: distributions of in-app and background values. Lower right: distribution of the within-patient, within-week difference.

3.3 Patients with large differences between the two gaits

The Bland–Altman plots show, for every parameter, a number of outliers whose in-app and background values diverge dramatically. Applying parameter-specific cutoffs, between 5 and 32 patients per parameter had at least one week with a very large difference (Table S1). Some showed a consistently large difference across most time points, others only in a small subset of weeks, and some differed in many parameters at once (Figure 3). Inspection of the underlying inertial patterns (Figure S4) showed that in some cases the two gaits were genuinely different, while in others the background recording did not resemble a valid gait pattern. No single explanation covered all cases.

A Large in-app–background differences in many parameters **B Similar in-app and background values**

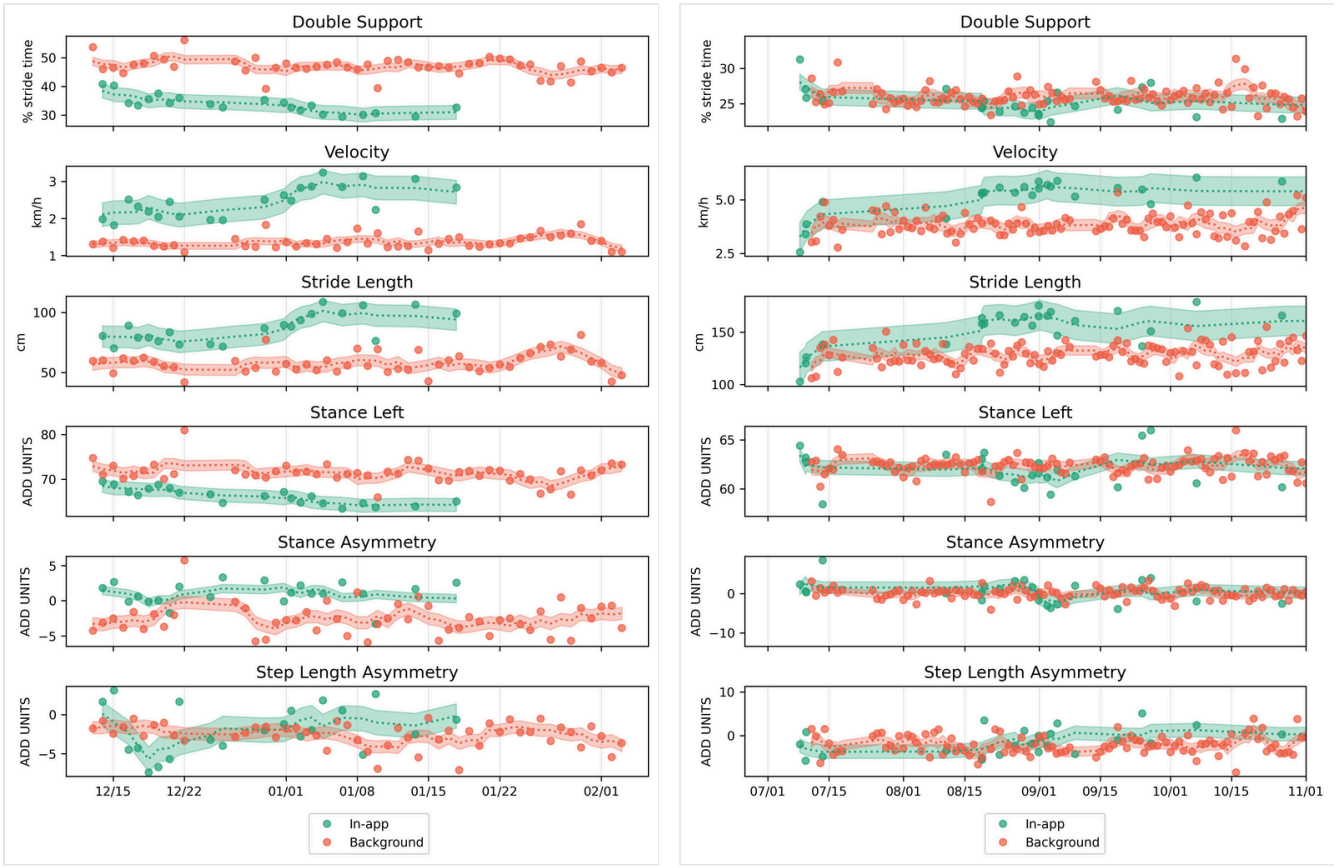


Figure 3. Two patients at the extremes of in-app–background agreement. A: a patient with large differences in many parameters. B: a patient with similar values; even here, velocity and stride length differ. Dots are individual in-app measurements and daily background aggregates; dotted lines are rolling averages with shaded standard deviation.

3.4 Recovery after joint replacement

Background walks tracked recovery after surgery in both THR (n = 45) and TKR (n = 32) patients (Figures 4 and 5), showing the typical pattern of rapid early recovery followed by slower improvement until parameters reach or slightly exceed pre-surgery values. The rate of recovery, however, differed between recording modes: background walks took noticeably longer than in-app walks to return to pre-surgery levels after both procedures.

© 2026 OneStep · www.onestep.co

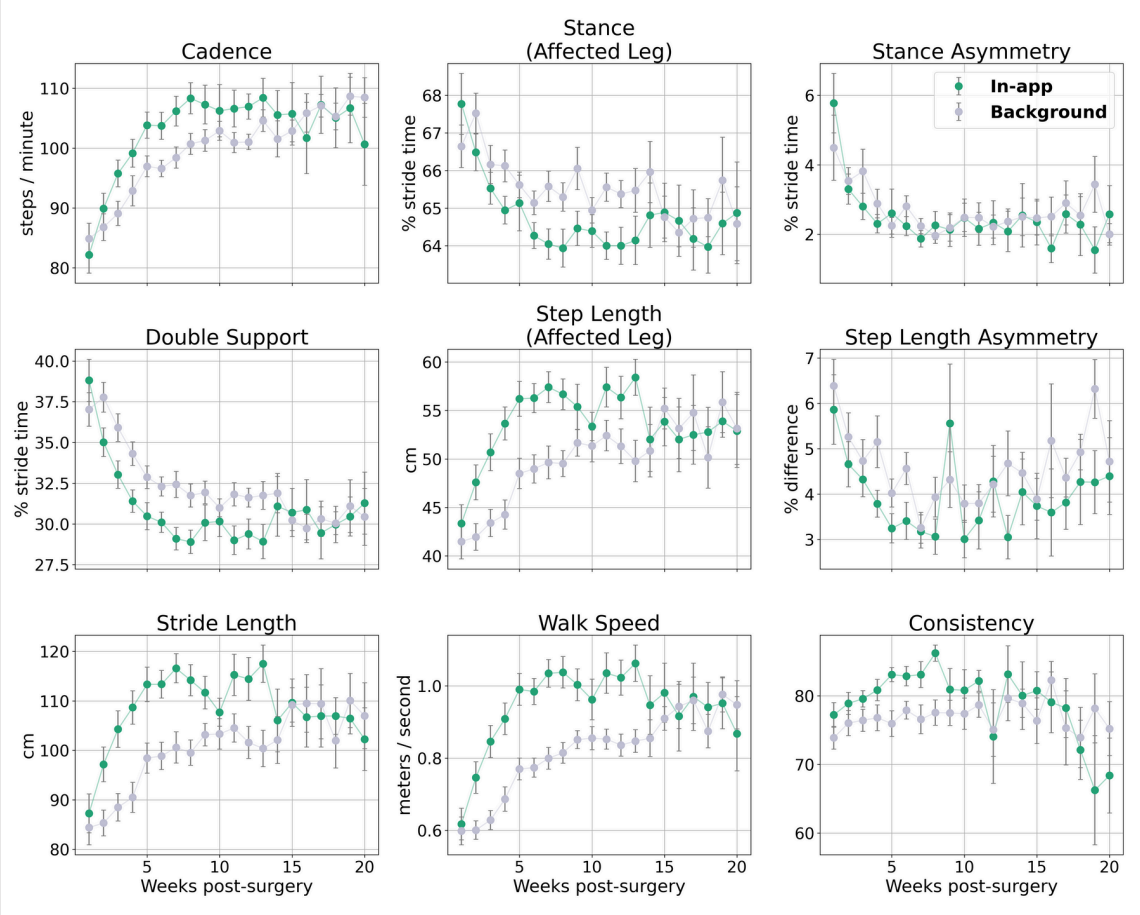


Figure 4. Recovery in in-app versus background walks after total hip replacement (n = 45). Points are the mean of weekly patient values; error bars are the standard error of the mean.

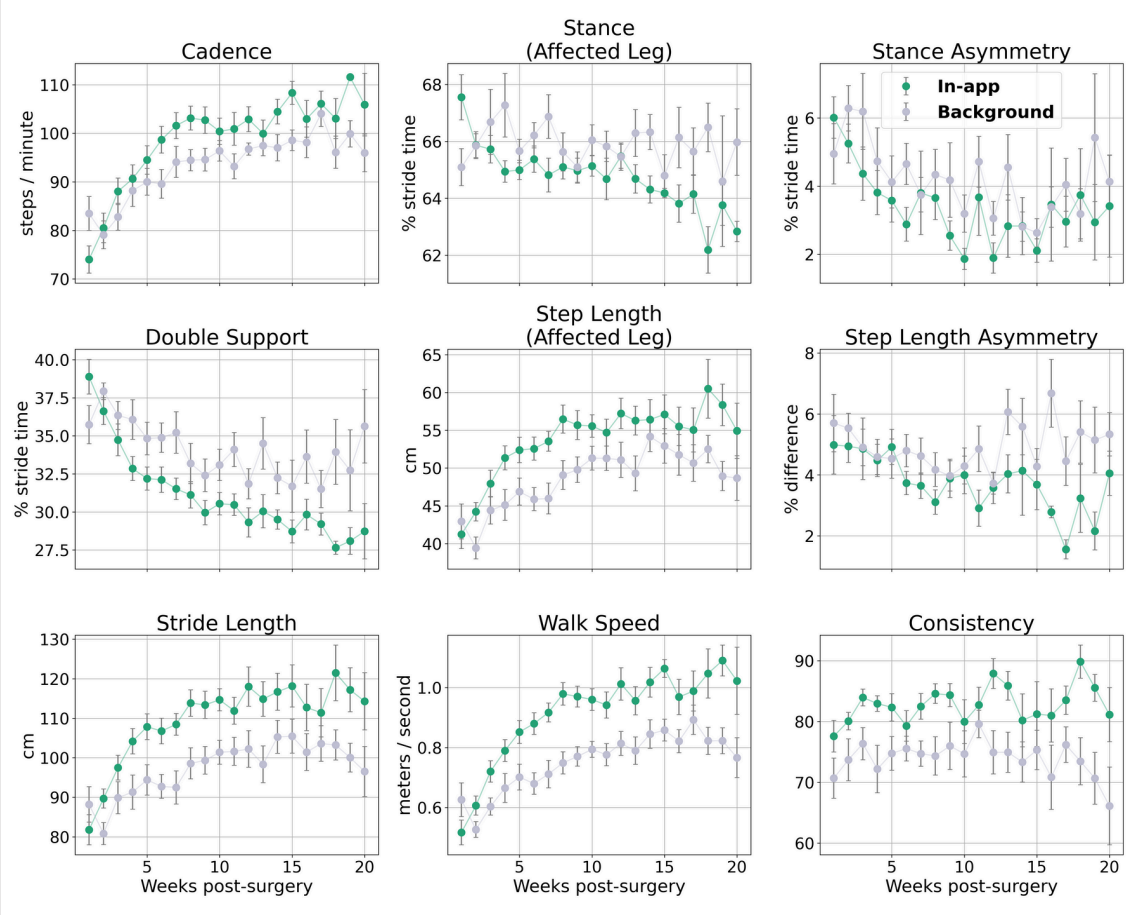


Figure 5. Recovery in in-app versus background walks after total knee replacement (n = 32). Points are the mean of weekly patient values; error bars are the standard error of the mean.

3.5 Detecting change in gait status

Short-term change. Comparing consecutive two-week periods, weeks in which both sources detected a significant change ($p < 0.05$) were enriched, but many changes were detected by only one source (Figure S5). In-app and background detections were significantly associated for short-term improvement ($\chi^2 = 38.3$, $p = 6.1 \times 10^{-10}$; Table S2) but not for short-term decline ($\chi^2 = 1.6$, $p = 0.21$; Table S3).

Change from baseline. Comparing each two-week period with the patient's baseline period, both improvement and decline showed strong association between sources (Figure 6; Tables 3 and 4). Among patients with alerts from both sources, the large majority had them in the same time period (Table S4).

Coverage. Background data nearly doubled the number of two-week periods with sufficient data for a change-of-status calculation: 2,184 versus 1,102 with in-app data (+98.2%). When in-app data were insufficient, background data detected 156 declines from baseline in 59 patients, 39 of whom never had a change of status detected from in-app data, and 482 improvements in 116 patients, 91 of whom had no in-app detection at any time.

© 2026 OneStep · www.onestep.co

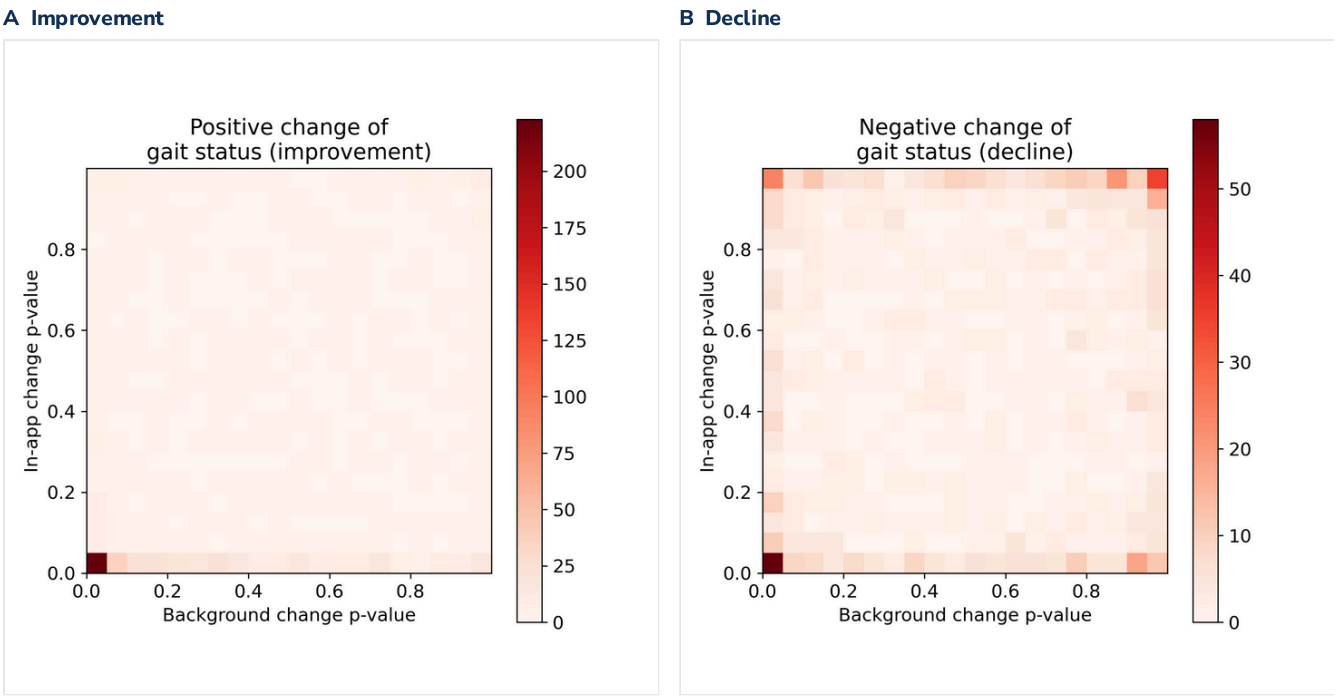


Figure 6. Change of status from baseline in in-app versus background gait. For each patient, a baseline two-week period was compared with every subsequent two-week period, yielding a p-value for change of status from each data source. Heatmaps show the joint distribution of in-app and background p-values for improvement (A) and decline (B).

Table 3. Positive change of status from baseline. $\chi^2 = 101.7$, $p = 6.52 \times 10^{-24}$.

	Significant in in-app	Not significant in in-app
Significant in background	223	73
Not significant in background	295	432

Table 4. Negative change of status from baseline. $\chi^2 = 25.0$, $p = 5.60 \times 10^{-7}$.

	Significant in in-app	Not significant in in-app
Significant in background	58	117
Not significant in background	133	669

4 Discussion

Background gait data, recorded when patients are not actively capturing their gait, offer a valuable view of mobility but are noisy, reflecting varied surfaces, footwear and conditions. With parameter-specific daily aggregation, this noise can be reduced enough for background data to be used clinically. After aggregation, most parameters agreed with in-app measurement at a good level of reliability, and the remaining difference was systematic rather than random: background gait is slightly less favorable on average, consistent with earlier work[7] and with the intuition that a conscious, in-app walk is a person's "best" walk while background walks span the range of walks a person actually takes. Unhealthy patterns that are suppressed under observation can emerge in background data.

The patients whose two gaits differ substantially are of particular clinical interest. Some may be able to control their gait when conscious of walking or being recorded, yet revert to unhealthy patterns when walking unconsciously. Identifying such patients matters for understanding their pathology and recovery, and background analysis offers a tool for doing so. In other cases the divergence reflected background recordings that were not valid walks, pointing to the need for further context-aware filtering.

Post-surgical recovery illustrates the practical value of passive data. Background walks returned to pre-surgery levels more slowly than in-app walks, implying that true recovery often takes longer than a clinician would conclude from clinic visits or prompted measurements alone. Background data allow clinicians to follow patients through the whole course of recovery and to know when a patient is genuinely back to normal.

© 2026 OneStep · www.onestep.co

A larger dataset may reveal how this gap varies between individuals.

Change-of-status alerts derived from background data create an opportunity to detect clinically relevant change throughout care without relying on patients to measure themselves: 156 declines and 482 improvements were detected when no in-app data were available. Short-term decline was not significantly associated between sources, possibly because such events are infrequent in these populations and obscured by noise after acute events such as surgery were excluded; change from baseline, which was highly significant for both directions, may be the more fruitful target. Because many changes were detected by only one source, alerts based on a single source should be treated with caution. We propose a two-tier scheme: a yellow alert when a change is detected from one source, and a red alert only when in-app and background changes coincide or other corroborating evidence exists.

Looking ahead, more sophisticated noise reduction, using the shape of the repetitive inertial pattern to infer the context of each walk, could separate background walks into categories such as walking while carrying an object, hurrying or distracted, enabling highly personalized analysis. Beyond its clinical uses, background measurement collects far more data with no added burden on patients, improving the accuracy of clinical indicators and deepening understanding of how people walk when conscious versus unconscious of their movement.

4.1 Limitations

The analysis is retrospective and drawn from patients using OneStep in rehabilitation. Background data remain noisier than in-app data, and some outlying background recordings may not be valid walks. Two parameters (consistency and step length asymmetry) agreed poorly between sources. Short-term decline detection did not agree between sources. The causes of large in-app–background divergence in individual patients were not fully explained.

5 Conclusion

With appropriate aggregation, background gait data agree well with in-app data for most parameters while capturing the slightly less healthy, everyday walk that conscious recording does not. Background data track recovery after joint replacement, reveal that real-world recovery lags prompted measurements, nearly double monitoring coverage and detect change when no in-app data exist. Passive gait monitoring is a feasible and valuable complement to conscious measurement, and the differences between the two are themselves a source of clinical insight.

References

1. Wren TA, Gorton GE 3rd, Õunpuu S, Tucker CA. Efficacy of clinical gait analysis: a systematic review. *Gait Posture.* 2011;34(2):149–153.
2. Ewen AM, Stewart S, St Clair Gibson A, Kashyap SN, Caplan N. Post-operative gait analysis in total hip replacement patients: a review of current literature and meta-analysis. *Gait Posture.* 2012;36(1):1–6.
3. Kwon MS, Kwon YR, Park YS, Kim JW. Comparison of gait patterns in elderly fallers and non-fallers. *Technol Health Care.* 2018;26(S1):427–436.
4. Muro-de-la-Herran A, Garcia-Zapirain B, Mendez-Zorrilla A. Gait analysis methods: an overview of wearable and non-wearable systems, highlighting clinical applications. *Sensors.* 2014;14(2):3362–3394.
5. Malone LA, Bastian AJ. Thinking about walking: effects of conscious correction versus distraction on locomotor adaptation. *J Neurophysiol.* 2010;103(4):1954–1962.
6. Warmerdam E, Hausdorff JM, Atrsaei A, et al. Long-term unsupervised mobility assessment in movement disorders. *Lancet Neurol.* 2020;19(5):462–470.
7. Rozanski G, Putrino D. Recording context matters: differences in gait parameters collected by the OneStep smartphone application. *Clin Biomech.* 2022;99:105755.
8. Stone E, Skubic M, Rantz M, Abbott C, Miller S. Average in-home gait speed: investigation of a new metric for mobility and fall risk assessment of elders. *Gait Posture.* 2015;41(1):57–62.
9. Rens N, Gandhi N, Mak J, et al. Activity data from wearables as an indicator of functional capacity in patients with cardiovascular disease. *PLoS One.* 2021;16(3):e0247834.
10. Picerno P, Iosa M, D'Souza C, Benedetti MG, Paolucci S, Morone G. Wearable inertial sensors for human movement analysis: a five-year update. *Expert Rev Med Devices.* 2021;18(sup1):79–94.
11. Christensen JC, Stanley EC, Oro EG, et al. The validity and reliability of the OneStep smartphone application under various gait conditions in healthy adults with feasibility in clinical practice. *J Orthop Surg Res.* 2022;17:417.
12. Shahar RT, Agmon M. Gait analysis using accelerometry data from a single smartphone: agreement and consistency between a smartphone application and gold-standard gait analysis system. *Sensors.* 2021;21:7497.
13. Shema-Shiratzky S, Beer Y, Mor A, Elbaz A. Smartphone-based inertial sensors technology: validation of a new application to measure spatiotemporal gait metrics. *Gait Posture.* 2022;93:102–106.
14. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med.* 2016;15(2):155–163.

© 2026 OneStep · www.onestep.co

Suggested citation. OneStep Research Team. How humans walk differently when conscious vs unconscious of their movement. OneStep White Paper; August 2023. Available at onestep.co/resources.

© 2026 OneStep · www.onestep.co

Supplementary figures and tables

Additional analyses supporting the main text. Figure S3 (per-patient double-support traces for outlier patients) is available in the original full-length report.

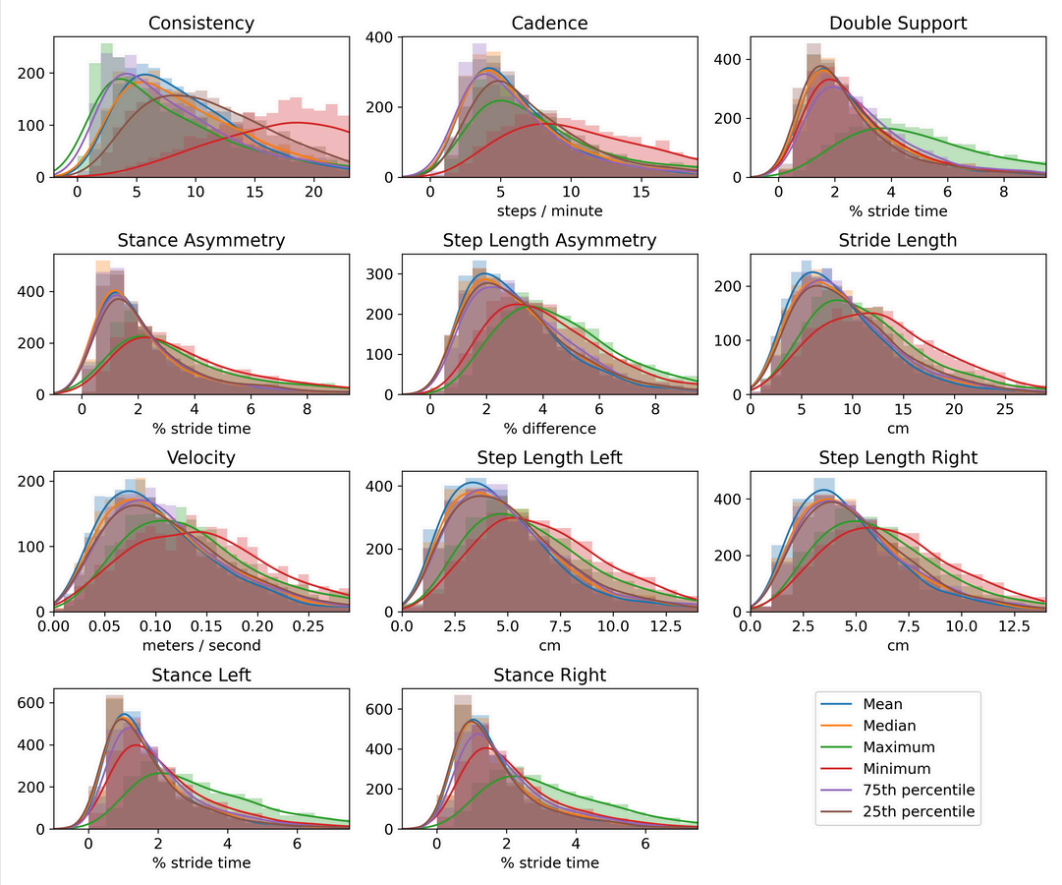


Figure S1. Full distribution of weekly standard deviation of aggregation methods. For each parameter, histograms of the weekly standard deviation of daily aggregates computed by six methods. The histogram shifted furthest left identifies the method that minimizes weekly variance.

© 2026 OneStep · www.onestep.co

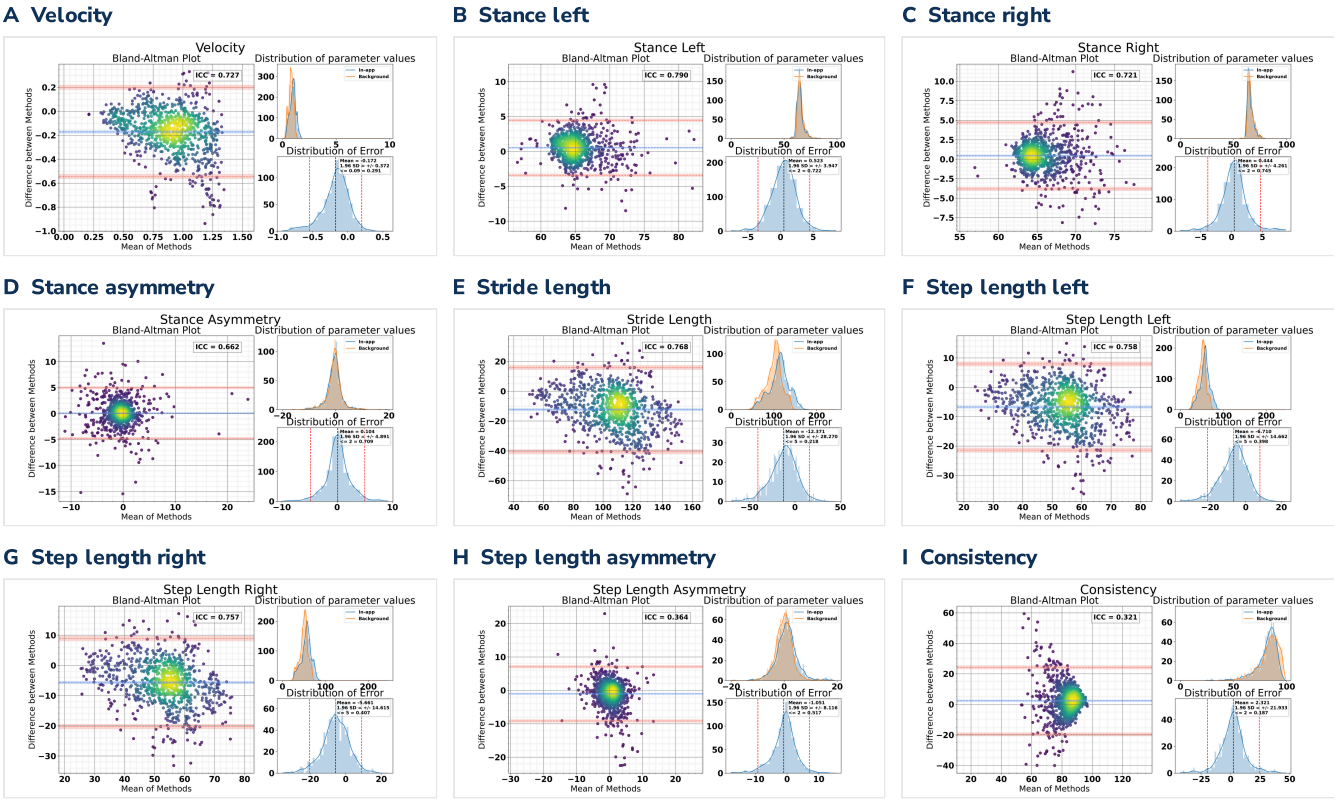


Figure S2. Comparison of weekly in-app and background values for the remaining nine parameters. Layout as in Figure 2: Bland–Altman plot with $ICC_{3,1}$ (left), distributions of in-app and background values (upper right) and of their difference (lower right).

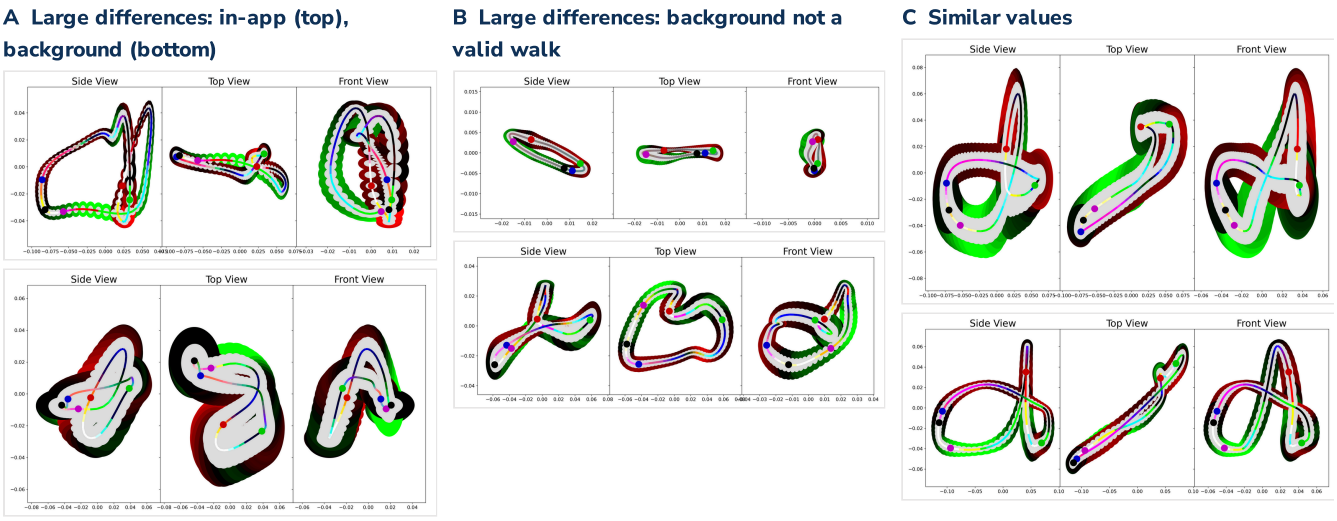


Figure S4. In-app versus background gait pattern reconstructions. Repetitive inertial pattern shapes (side, top and front views) for in-app (top) and background (bottom) walks. A: a patient with very different parameter values between modes; the shapes differ markedly. B: another such patient, in whom the background pattern does not resemble a legitimate gait pattern. C: a patient with similar values. The central line traces movement, colored by direction; the aura width is the standard deviation of acceleration. The black dot marks maximum hip extension; red, green, blue and purple dots mark primary heel strike, secondary toe-off, secondary heel strike and primary toe-off.

© 2026 OneStep · www.onestep.co

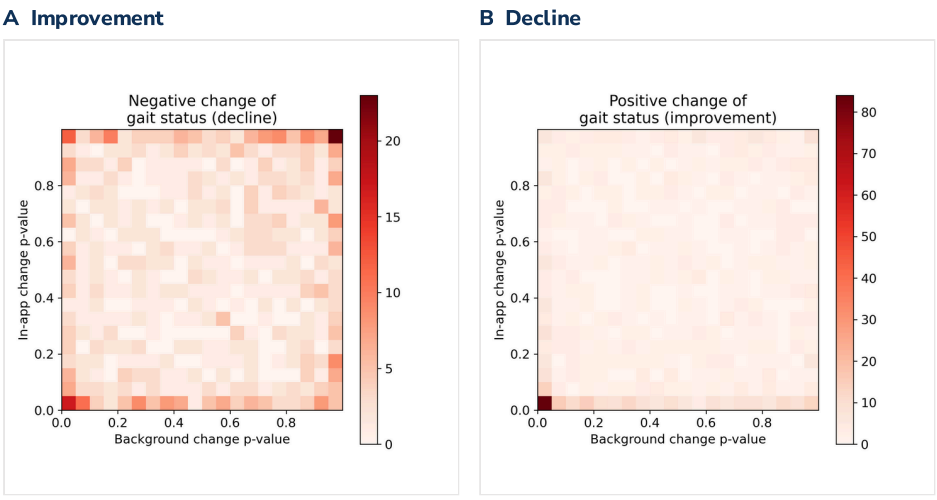


Figure S5. Short-term (acute) change of status in in-app versus background gait. For every pair of consecutive two-week periods, a p-value for change of status was computed from each source. Heatmaps show the joint distribution for improvement (A) and decline (B).

Table S1. Patients with very large in-app–background differences, by parameter and cutoff.

Parameter	Cutoff	Patients
Double support	10% stride time	9
Gait speed	0.56 m/s	7
Stride length	30 cm	32
Left stance	5% stride time	18
Stance asymmetry	10% stride length	5
Step length asymmetry	10% stride time	15

Table S4. Timing of in-app (IA) and background (BG) change-of-status-from-baseline alerts (number of patients).

	No alerts	BG only	IA only	Both, different time	Both, same time
Improvement	118	46	46	3	30
Decline	127	40	43	3	27

Table S2. Positive short-term change of status. $\chi^2 = 38.3$, $p = 6.14 \times 10^{-10}$.

	Significant in in-app	Not significant in in-app
Significant in background	84	91
Not significant in background	178	552

Table S3. Negative short-term change of status. $\chi^2 = 1.59$, $p = 0.207$.

	Significant in in-app	Not significant in in-app
Significant in background	17	81
Not significant in background	102	697

© 2026 OneStep · www.onestep.co