OneStep

WHITE PAPER · FALL RISK ASSESSMENT

Comparing the OneStep Fall Risk Assessment With Standard Clinical Measures

Convergent validity and fall-prediction accuracy against the Berg Balance Scale, Timed Up & Go and sit-to-stand

ONESTEP RESEARCH | OCTOBER 2026 | ONESTEP.CO/RESOURCES

RE Comparing the OneStep Fall Risk Assessment With Standard Clinical Measures: Convergent validity and fall-prediction accuracy against the Berg Balance Scale, Timed Up & Go and sit-to-stand

One score that agrees with the tests you trust, predicts falls better, and takes a minute

Clinicians judge fall risk with the Berg Balance Scale, the Timed Up & Go and the sit-to-stand. Each is validated, each takes clinician time, and each measures one facet of a multi-dimensional problem. This paper compares the OneStep Fall Risk Assessment (FRA) with those standard measures on two counts: does it measure the same thing, and does it predict real falls better?

r = 0.82
Agreement between the FRA and the Timed Up & Go (n = 17,681)

0.74
C-index of the FRA for 25-day falls, same residents

0.56 / 0.61
C-index of TUG and sit-to-stand used alone

~1 min
To administer, by any trained staff member

On both counts the answer favors the FRA. It correlates with every standard measure in the expected direction (TUG +0.82, Berg −0.71, sit-to-stand −0.68, Tinetti −0.58), so it captures the balance-and-mobility construct those scales measure. And in the same residents, measured against prospectively recorded falls, it ranked who would fall with a C-index of 0.74, against 0.56 for the Timed Up & Go and 0.61 for sit-to-stand, both differences significant. The two bedside tests were no better than each other.

Potential to replace

A score that reproduces what the standard tests measure, and predicts falls better, can stand in for them in routine screening.

Time back to the clinician

A Berg takes 15–20 minutes of a therapist's time; the FRA takes a one-minute walk with a phone in the pocket, run by any trained staff member.

Keep the tests where they add depth

Where a Berg or TUG is still performed, it feeds the FRA and sharpens the score. Nothing clinicians trust is discarded.

PART OF THE VALIDATION PACKET

This paper is one of three. **Clinical Validation of the OneStep Fall Risk Assessment** shows that higher FRA tiers fall sooner and more often. This paper shows the FRA agrees with the standard measures and out-predicts them. **One Walk Is Enough** shows that a single one-minute walk already carries most of that accuracy.

© 2026 OneStep

The full white paper follows.

WHITE PAPER · OCTOBER 2026

Comparing the OneStep Fall Risk Assessment With Standard Clinical Measures

Convergent validity and fall-prediction accuracy against the Berg Balance Scale, Timed Up & Go and sit-to-stand

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

ABSTRACT

Background. Bedside tests such as the Berg Balance Scale, Timed Up & Go (TUG) and sit-to-stand (STS) are validated and widely used, but each samples one facet of fall risk at one moment, and standalone functional tests typically show only moderate predictive accuracy. The OneStep Fall Risk Assessment (FRA) integrates these tests with smartphone-measured gait, everyday walking and clinical history into one continuous score.

Objective. To establish that the FRA measures the same construct as established tools, and to test whether the integrated score predicts falls better than single bedside tests in the same residents.

Methods. Residents of senior-living and post-acute communities across four partner organizations with prospectively logged fall incidents. Convergent validity was assessed by Spearman correlation between the FRA and TUG, STS, Berg and Tinetti. Discrimination was compared among residents with an FRA, TUG and STS at baseline ($n \approx 366$) by the concordance index (C-index) for time to first fall over 25 days, with differences tested by paired bootstrap on the same residents.

Results. The FRA correlated with every established test in the expected direction: TUG $r = +0.82$ ($n = 17{,}681$), STS $r = -0.68$, Berg $r = -0.71$, Tinetti $r = -0.58$. Berg and TUG correlated only moderately with each other ($r = -0.64$). The integrated FRA reached a C-index of 0.74, versus 0.56 for TUG alone (0.17 lower, $p < 0.001$) and 0.61 for STS alone (0.13 lower, $p = 0.002$); TUG and STS did not differ from each other.

Conclusions. The FRA reflects the signal clinicians already trust and adds to it. In the same residents, measured against real falls, the integrated score predicts who will fall substantially better than any single bedside test.

KEY POINTS

Question	Does combining bedside tests with smartphone-measured gait and history predict falls better than any single test used alone?
Findings	In the same residents, the integrated FRA predicted who would fall with a C-index of 0.74, versus 0.56 for the Timed Up & Go and 0.61 for sit-to-stand (both differences significant); the two single tests were no different from each other. The FRA correlated with TUG, STS, Berg and Tinetti in the expected direction.
Meaning	Clinicians need not abandon the tests they trust; integrating them with gait and history gives a materially more accurate ranking of who is at risk.

KEYWORDS fall risk · Berg Balance Scale · Timed Up and Go · sit-to-stand · concordance index · convergent validity · smartphone gait analysis · senior living · post-acute care

1 Introduction

A fall is rarely the result of one deficit. It emerges from the interplay of balance, strength, gait, endurance, cognition, fear and history. A bedside test, by design, samples one of these at one moment in a clinic: the Berg Balance Scale captures standing balance, the Timed Up & Go (TUG) mobility and transfers, the sit-to-stand (STS) lower-limb strength. That is why even well-validated tools show only moderate accuracy for predicting falls on their own, and why a clinician often reaches for two or three of them.

© 2026 OneStep · www.onestep.co

The OneStep Fall Risk Assessment (FRA) does not replace these measures; it integrates them. At its core is a short one-minute walk, recorded with a smartphone in the resident's pocket, that measures gait in detail. The FRA combines that gait with functional tests, turning, everyday walking and clinical history into one continuous score, validated against time to first fall in the companion paper.[1] This paper asks whether the integration is worth it: does the combined score predict falls better than acting on any one test alone?

2 Methods

2.1 Cohort and outcome

Residents of senior-living and post-acute communities across four partner organizations in which fall incidents are prospectively logged as structured electronic health record data. The FRA was confidence-screened (v4.0.0). For the head-to-head comparison, residents with an FRA, a TUG and an STS were included, each test taken as the nearest measurement to baseline ($n \approx 366$; the same raw-test pairing used in the companion one-minute-walk paper[2]). The outcome was time to each resident's first fall from structured incident records, with single-spell, discharge-aware censoring.

2.2 Convergent validity

Agreement between the FRA and the established tools was quantified by Spearman correlation on paired measurements. For TUG, each FRA was paired with the nearest TUG within 90 days (closest pair per resident); for the Berg-versus-TUG comparison, residents assessed with both were used ($n = 212$).

2.3 Discrimination

Discrimination was measured by the concordance index at a 25-day horizon. Differences between predictors were tested by paired bootstrap on the same residents, so each comparison is like-for-like. The analyses treat the FRA as a fixed, pre-existing model and the falls as an external outcome it was not trained on.

2.4 Scope and considerations

TUG and STS are themselves inputs to the FRA, which is stated plainly here; the comparison reflects integrated-score versus single-test practice rather than an in-sample fit. Because the head-to-head uses the subset with all three measures, the FRA's absolute accuracy here (0.74) runs above the full-cohort figure (about 0.64 in the companion papers); the claim is the paired difference between predictors on identical residents, not the absolute value. The Berg and Tinetti scales were too sparse in the falls cohort for a direct head-to-head and are reported on convergent-validity grounds; their fall-prediction comparison is a target for future data. All data were de-identified.

WHAT IS THE C-INDEX?

The concordance index is a plain measure of how well a score ranks risk. Picture two residents, one who later fell and one who did not: the C-index is how often the score gave the higher risk to the one who actually fell. 0.50 is a coin toss and 1.00 is perfect. In fall-risk research, single bedside tests usually land around 0.60.

3 Results

3.1 Each single test sees only part of the picture

If the common tests truly captured the same underlying risk, they would agree almost perfectly. They do not. Across residents assessed with both, the Berg Balance Scale and the TUG correlate only moderately (Spearman $r = -0.64$; Figure 1). They share roughly 40% of their variation and disagree on the rest. Each carries information the other misses.

© 2026 OneStep · www.onestep.co

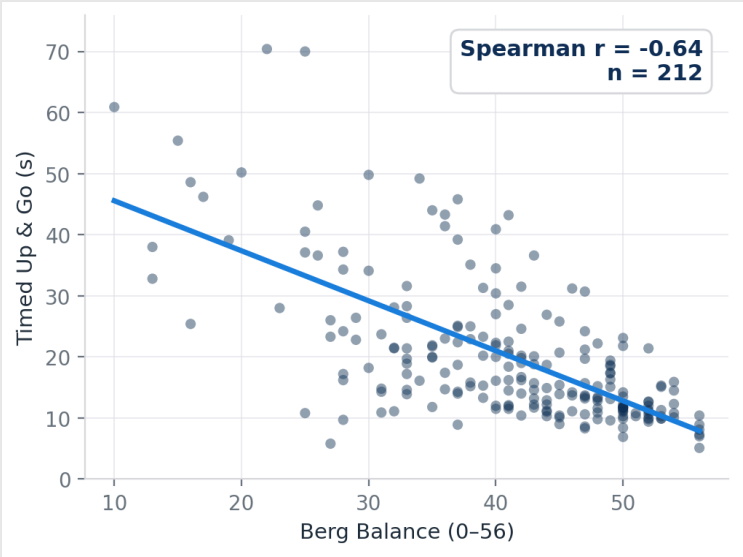


Figure 1. Two trusted tests, only partial agreement. Berg Balance Scale versus Timed Up & Go in the same residents (n = 212; Spearman r = −0.64). Better balance tends to mean a faster TUG, but the scatter is wide, and neither test substitutes for the other.

3.2 The FRA speaks the same clinical language

Before asking whether the FRA predicts falls better, it is worth confirming that it measures the right thing. It does. The FRA agrees with every one of the established tests, each in the expected direction: TUG r = +0.82 (n = 17,681 paired measurements; Figure 2), sit-to-stand r = −0.68 (n = 20,997), Berg r = −0.71 (n = 299) and Tinetti r = −0.58 (n = 133) (Figure 3). Because TUG time is bounded below and long-tailed above, the relationship is linear on a logarithmic TUG scale (Pearson r on log TUG = +0.80). The FRA does not discard the signal clinicians trust; it reflects it, and then adds to it.

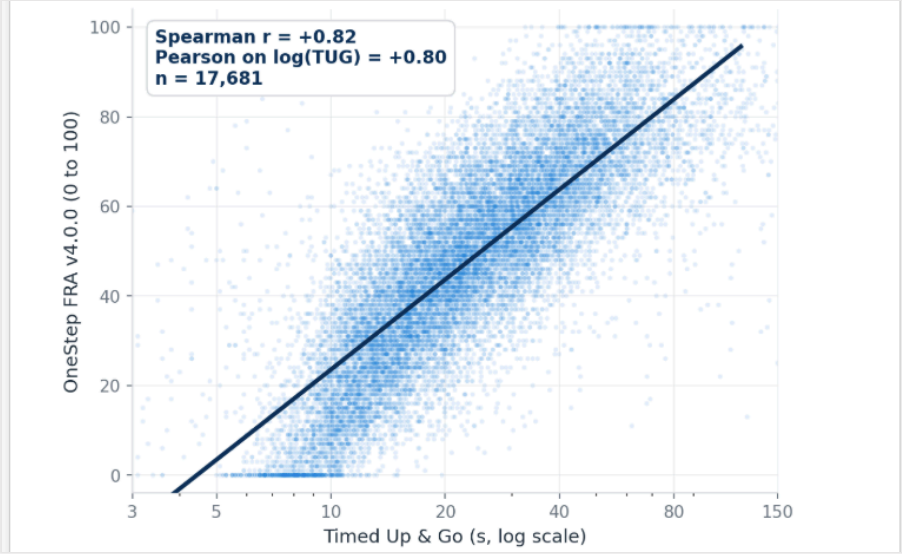


Figure 2. The FRA aligns with the Timed Up & Go. OneStep FRA v4.0.0 against TUG time on a logarithmic scale, each point one resident (FRA paired with the nearest TUG within 90 days, closest pair per resident; n = 17,681). Spearman r = +0.82; Pearson r on log(TUG) = +0.80. Slower TUG, higher FRA risk.

© 2026 OneStep · www.onestep.co

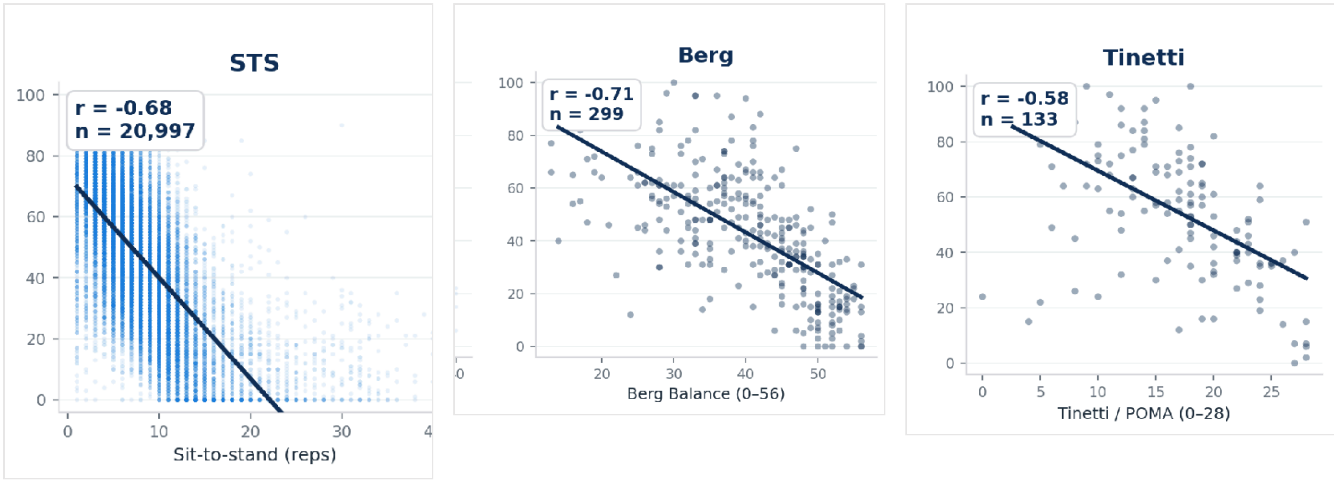


Figure 3. The FRA aligns with sit-to-stand, the Berg Balance Scale and the Tinetti (POMA). OneStep FRA v4.0.0 (0–100) against each test, one point per resident; r = Spearman correlation. More repetitions, higher Berg or higher Tinetti all correspond to lower FRA risk, as expected.

3.3 Head to head: the integrated score predicts real falls better

The decisive test is predictive, not correlational: in the same residents, how accurately does each approach rank who actually falls? Among residents with all three measures at baseline, the integrated FRA reached a C-index of 0.74, against 0.56 for TUG alone and 0.61 for STS alone (Figure 4, Table 1). Both differences were significant ($p < 0.001$ and $p = 0.002$). Just as telling, TUG and STS were no different from each other: swapping one single test for another buys nothing. Only the combination moves the needle. The single tests cluster at the moderate accuracy the literature reports for standalone functional tests; the integrated score sits well above them.

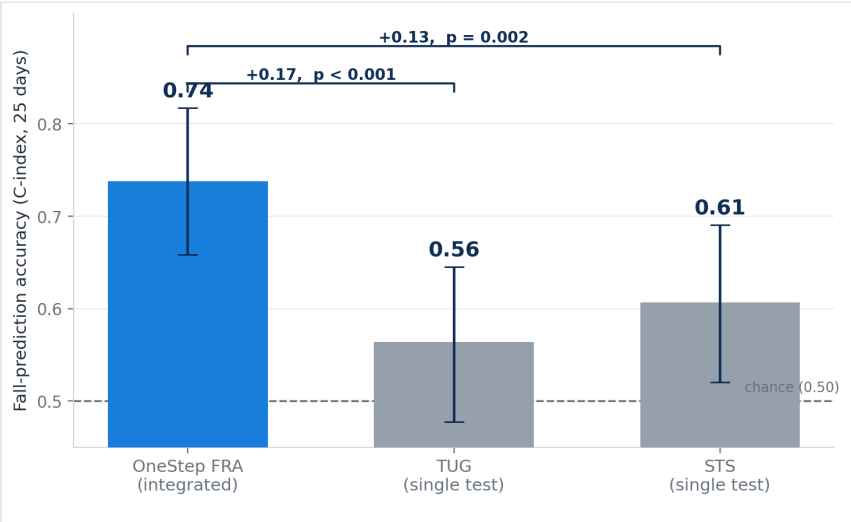


Figure 4. Fall-prediction accuracy (concordance index) in the same residents. The integrated FRA (0.74) significantly out-ranks both single tests (TUG 0.56, STS 0.61), which are statistically indistinguishable from each other. Bars show 95% bootstrap intervals; the dashed line marks chance (0.50). Same residents (n ≈ 366), 48 first falls within 25 days, paired bootstrap.

Table 1. 25-day fall-prediction accuracy in the same residents (paired bootstrap).

Predictor	C-index	Versus the FRA
OneStep FRA (integrated)	0.74	Reference
Timed Up & Go, alone	0.56	0.17 lower ($p < 0.001$)
Sit-to-stand, alone	0.61	0.13 lower ($p = 0.002$)

© 2026 OneStep · www.onestep.co

4 Discussion

The FRA's advantage is not a statistical trick; it has more to work with. Beyond the bedside tests, it incorporates signals a one-time, in-clinic measurement cannot: a guided one-minute walk that measures how the resident actually moves (gait speed, cadence, double-support time, left–right symmetry and turning), everyday gait captured passively during normal activity, the resident's own fall history and fear of falling, and, because it is captured continuously, a trajectory rather than a snapshot. A clinician choosing to act on a single number is, in effect, discarding most of this. The FRA keeps it all and weighs it together.

There is a further, practical distinction. The fall-prediction evidence for the Berg and TUG comes from the published literature, in other populations. The FRA has been validated against observed falls in these residents directly: in the companion validation, residents it rated Very High fell at roughly one in four within 25 days, while those it rated Low had none.[1] The score a clinician sees is the score that was tested against what actually happened.

5 Conclusion

The OneStep FRA agrees with the Berg, TUG and sit-to-stand, integrates them with a one-minute walk, everyday gait and history, and, measured against real falls in senior-living and post-acute residents, predicts who falls more accurately than any of those tests used on its own. Keep the tests you trust; add the picture they miss.

References

1. OneStep Research Team. Clinical validation of the OneStep Fall Risk Assessment: time-to-event stratification in senior living and post-acute care. OneStep White Paper; September 2026. Available at onestep.co/resources.
2. OneStep Research Team. One walk is enough: a single one-minute walk gives a fall-risk score as useful as the full assessment. OneStep White Paper; October 2026. Available at onestep.co/resources.

Suggested citation. OneStep Research Team. Comparing the OneStep Fall Risk Assessment with standard clinical measures: convergent validity and fall-prediction accuracy against the Berg Balance Scale, Timed Up & Go and sit-to-stand. OneStep White Paper; October 2026. Available at onestep.co/resources.

© 2026 OneStep · www.onestep.co