RESEARCH | Article in Press

Timed Up and Go, gait speed, and grip strength: associations and discriminative performance for falls in community-dwelling older adults - results from the KORA-age study

Sebastian Martini · Michael Drey · Ralf Schmidmaier · Uta Ferrari · Canan Marcon · Annette Peters · Marie-Theres Huemer · Margit Heier · Barbara Thorand

Received: 21 July 2026 / Accepted: 28 October 2026
© The Author(s) 2026

Abstract

Background Physical performance measures are commonly used in geriatric assessment and have been associated with falls, but their ability to discriminate between older adults with and without falls may be limited. We compared the Timed Up and Go (TUG) test, gait speed, and grip strength within a population-based cohort, considering associations, discriminatory performance, incremental discrimination, and established clinical TUG thresholds.

Methods We analyzed data from 1,079 community-dwelling adults aged≥65 years participating in KORA-Age. Fall status was classified for 819 participants based on self-reported falls in the 12 months preceding the 3-year follow-up interview. Associations of baseline TUG, gait speed, and grip strength with reported falls were examined using sequential logistic regression models adjusted for age and sex, additional clinical covariates, and frailty. Discriminatory performance was assessed using the area under the receiver operating characteristic curve (AUC), with DeLong tests used to evaluate incremental discrimination beyond age and sex. We additionally examined combinations of the three performance measures and predefined TUG thresholds. Inverse-probability weighting was used in sensitivity analyses to address selective follow-up.

Results Poorer TUG performance remained associated with reported falls after additional adjustment for frailty (OR per 1-s increase 1.095, 95% CI 1.018–1.176), whereas the association with slower gait speed was attenuated after frailty adjustment and grip strength was not associated with falls. The age- and sex-adjusted reference model had an AUC of 0.680; corresponding AUCs after adding TUG, gait speed, or grip strength were 0.700, 0.692, and 0.675, respectively, with no clear evidence of incremental discrimination. Combining all three performance measures did not improve discrimination beyond TUG alone. Higher TUG thresholds identified smaller groups with a higher prevalence of falls and greater specificity but progressively lower sensitivity. Inverse-probability-weighted analyses yielded similar associations.

Conclusion TUG remained associated with reported falls after multivariable and additional frailty adjustment, but none of the three physical performance measures showed clear incremental discrimination beyond age and sex. Combining the measures did not improve discrimination, and established TUG thresholds showed limited

This Article in Press is shared early to give you faster access to new research. It is citable and carries a permanent DOI. The final edited version will replace it automatically.

Published online: 06 November 2026

sensitivity. These findings support distinguishing statistical associations with falls from the ability of physical performance measures to discriminate between older adults with and without falls.
Clinical trial number Not applicable.

Keywords Falls · Fall risk assessment · Geriatric assessment · Timed Up and Go · Gait speed · Handgrip strength

Background

Falls are common in older adults and are associated with injuries, loss of independence, reduced quality of life, hospitalization, and mortality [1]. Accordingly, international guidelines emphasize early identification of individuals at increased risk and recommend multifactorial assessment and prevention strategies in community-dwelling older adults [2–4].

Physical performance measures are commonly used in geriatric assessment because they are simple, inexpensive, and feasible in both clinical and epidemiological settings. The Timed Up and Go (TUG) test, gait speed, and grip strength capture partly distinct aspects of physical function: gait speed reflects walking performance, grip strength provides a measure of muscle strength, whereas TUG integrates several components of mobility, including standing up, walking, turning, and sitting down [5, 6]. Their association with falls has been examined extensively, but an association with falls does not necessarily imply an ability to discriminate reliably between individuals with and without falls.

Previous studies have directly compared TUG, walking speed, and grip strength and generally found only modest discrimination for any falls, while stronger performance has been reported for recurrent falls [7, 8]. Combining several physical performance measures has likewise not consistently improved discrimination for any falls [8]. In KORA-Age, grip strength has previously been associated with falls in an age- and sex-adjusted analysis, but TUG and gait speed were not evaluated within the same analytical framework [5]. Thus, evidence remains limited on how these commonly used measures compare in a population-based cohort when associations, discriminatory performance, incremental discrimination, and clinically established thresholds are considered separately.

We therefore compared baseline TUG, gait speed, and grip strength with regard to their associations with reported falls in the 12 months preceding the 3-year follow-up interview in the population-based KORA-Age study. We further evaluated their discriminatory performance and incremental discrimination beyond age and sex, examined whether combining the three measures improved discrimination, and explored the classification characteristics of established clinical TUG thresholds. In an additional exploratory analysis, we also examined associations and discriminatory performance for recurrent falls.

Methods

Study population

The data for this study were obtained from the Cooperative Health Research in the Region of Augsburg (KORA)-Age study, a study including individuals aged 65 years and older who had previously participated in the Monitoring Trends and Determinants in Cardiovascular Diseases (MONICA)/KORA studies [9]. Detailed information on the study design, sampling procedures, and data collection has been published previously [10]. Initially, a questionnaire was sent to all eligible participants followed by a telephone interview. Finally, a sex and age-stratified sample was examined at the KORA-Age study center in 2008/2009. Of the 1,079 participants examined at the study center (baseline examination), 822 had follow-up information on the trigger question regarding dizziness,

balance problems, or problems with falling during the 12 months preceding the 3-year follow-up interview. Three participants responded "don't know" to the trigger question and were not assigned a fall status. Thus, classifiable follow-up fall status was available for 819 participants, whereas fall status could not be classified for 260 participants. Analyses were based on available data for the respective variables and models; therefore, sample sizes varied between analyses.

All participants provided written informed consent prior to inclusion, and all procedures were conducted in accordance with the Declaration of Helsinki.

Clinical assessment

Height and weight were measured using a standardized wall stadiometer (SECA 242, Seca, Hamburg, Germany) and a mechanical scale (SECA 635, Seca, Hamburg, Germany). Physical activity was assessed using the Physical Activity Scale for the Elderly (PASE), a validated questionnaire designed to capture occupational, household, and leisure-time activities over the previous week [11]. Comorbidity burden was defined as the number of 13 predefined chronic conditions, as described previously for the KORA-Age cohort [12]. Information on previous falls in the 12 months before baseline and medication use, including opioid analgesics, was collected via standardized telephone interviews and self-administered questionnaires. Serum 25-hydroxyvitamin D (25OHD) levels (ng/mL) were measured at baseline using an electrochemiluminescence immunoassay (ECLIA, Elecsys 2010, Roche Diagnostics GmbH, Mannheim, Germany) [13].

Gait speed

Gait speed was assessed in meters per second using the GAITRite system (CIR Systems, Havertown, PA, USA), which quantifies spatiotemporal gait parameters over a 488 × 61 cm walkway [14].

Grip strength

Grip strength was assessed using the JAMAR handheld dynamometer (Saehan Corp., Masan, Korea). Measurements were performed using the dominant hand. Participants completed three attempts with brief pauses between attempts and were encouraged to exert their maximum strength. The highest value of the three attempts was used for analysis.

Timed up and go test

The TUG test measures the time required to stand up from a chair, walk 3 m, turn, walk back, and sit down. Time was recorded in seconds using an electronic stopwatch.

Appendicular muscle mass and appendicular skeletal muscle index

Appendicular skeletal muscle mass (ASMM) was assessed using bioelectrical impedance analysis (BIA 2000-S; Data Input GmbH, Frankfurt, Germany). ASMM was calculated using the equation of Sergi et al. based on bioelectrical impedance analysis-derived parameters (BIA-2000-S Analyzer (Data Input, Frankfurt, Germany)) [15]. The appendicular skeletal muscle mass index was calculated as ASMM divided by height squared and expressed in kg/m^2.

Sarcopenia

Sarcopenia was defined according to the European Working Group on Sarcopenia in Older People (EWGSOP2) criteria with the three categories "no sarcopenia", "probable sarcopenia" and "confirmed sarcopenia" [16]. The

category "probable sarcopenia" includes participants with low handgrip strength (men < 27 kg, women < 16 kg) but normal appendicular skeletal muscle mass (ASMM, men ≥ 20 kg, women ≥ 15 kg). The category "confirmed sarcopenia" includes those with low handgrip strength and low ASMM (men < 20 kg, women < 15 kg).

Frailty status

Frailty was assessed according to the phenotype described by Fried et al. [17], including five components: unintentional weight loss, slow walking speed, low grip strength, low physical activity, and self-reported exhaustion. Participants meeting one or two criteria were classified as prefrail, and those meeting three or more criteria as frail.

Cognitive status

Cognitive status was assessed using the education-adjusted modified Telephone Interview for Cognitive Status (TICS-m), resulting in three categories: impaired cognition, mild cognitive impairment, and normal cognition [18].

Follow-up assessment and fall outcome

A follow-up assessment was conducted approximately 3 years after baseline using standardized telephone interviews. The KORA-Age study protocol allowed proxy interviews with relatives, friends or caregivers when participants were unable to complete an interview themselves because of physical or mental impairment [10]. At the Age2 follow-up, 29 of 822 interviews (3.5%) were conducted with proxies [19]. Participant-level information identifying proxy interviews was not available in the analytical dataset used for the present study.

Fall-related information was collected as part of the standardized KORA-Age assessment using questions that have been described previously [5]. The initial screening question concerning dizziness, balance problems, or problems with falling was adapted from item BAQ010 of the National Health and Nutrition Examination Survey (NHANES) 2003–2004 Balance Questionnaire [20]. All questions referred to the 12 months preceding the 3-year follow-up interview. Participants who answered the screening question affirmatively were subsequently asked a detailed fall question distinguishing between one fall, multiple falls, and no fall. According to the questionnaire routing, the detailed fall question was not administered to participants who denied the screening question.

For the primary fall outcome, participants reporting one or multiple falls in the detailed fall question were classified as fallers. Participants reporting no fall in the detailed question, as well as participants who denied the initial screening question, were classified as non-fallers. Participants with missing responses or an uncertain response to the screening question were not assigned a fall status. Thus, the outcome reflects reported falls at follow-up interview and does not represent continuous ascertainment of falls throughout the entire interval between baseline and follow-up.

Statistical analysis

Descriptive analyses

Baseline characteristics were summarized according to fall status. Continuous variables are presented as median and interquartile range, and categorical variables as n and percentage. Available N is shown for each characteristic. Standardized mean differences (SMDs) were used to quantify between-group differences in baseline characteristics, both between fallers and non-fallers and between participants with and without classifiable fall status at the 3-year follow-up. Although continuous variables are presented as medians and interquartile ranges, SMDs were calculated from the raw data as the difference in group means divided by the pooled standard deviation. For

categorical variables, a single overall SMD was calculated for each variable rather than separately for individual categories.

Association analyses

Associations between baseline physical performance measures and fall status at the 3-year follow-up were examined using logistic regression models. Three sequential logistic regression models were fitted for each physical performance measure. Model 1 was adjusted for age and sex. Model 2 additionally included previous falls, comorbidity burden, cognitive status, opioid analgesic use, and physical activity. Model 3 further included frailty status according to the Fried phenotype. Because the Fried phenotype incorporates slow walking speed, low grip strength, and low physical activity, Model 3 was considered an extended frailty-adjusted model and was interpreted cautiously for gait speed and grip strength because of overlap between these performance measures and the frailty construct. TUG was modeled per 1-s increase, gait speed per 0.1-m/s decrease, and grip strength per 1-kg decrease, such that odds ratios consistently represent worsening physical performance. Analyses were based on complete cases for the variables included in the respective model, and the number of participants and fall events is reported for each model.

Discrimination and exploratory analyses

Discriminative performance was evaluated using the area under the receiver operating characteristic curve (AUC) with 95% confidence intervals. To ensure direct comparability, the main ROC analyses were performed in a common complete-case sample including fall status, age, sex, TUG, gait speed, and grip strength. The age- and sex-adjusted model served as the reference model and was compared with models additionally including TUG, gait speed, grip strength, or all three physical performance measures. The age- and sex-adjusted TUG and gait-speed models were also compared directly. Differences between correlated AUCs were assessed using DeLong tests.

In exploratory incremental discrimination analyses, four nested models were evaluated in the same common complete-case sample. The reference model included age and sex. A clinical model additionally included previous falls in the 12 months before baseline and opioid analgesic use at baseline. TUG was subsequently added to the clinical model, followed by a model additionally including gait speed and grip strength. AUCs were compared sequentially using DeLong tests, with each model compared with the preceding nested model.

Exploratory analyses evaluated three predefined TUG thresholds (> 12 s, ≥ 13.5 s, and > 15 s) selected from previously published clinically relevant cut-points [21–23]. The > 12-s threshold was derived from a large population-based cohort of older adults in which fracture risk increased more markedly above this level [21]. The ≥ 13.5-s threshold originates from a frequently cited study evaluating TUG in relation to fall risk in community-dwelling older adults [23]. The > 15-s threshold was selected based on the 2022 World Guidelines for Falls Prevention and Management [22]. These predefined thresholds were examined to characterize the classification performance of previously proposed cut-points rather than to derive an optimal threshold from the present cohort. For each threshold, sensitivity, specificity, positive predictive value, negative predictive value, and age- and sex-adjusted odds ratios were calculated.

In a separate exploratory analysis, recurrent falls were defined as reporting multiple falls in the detailed fall question. All exploratory analyses were considered hypothesis-generating.

Attrition and sensitivity analyses

To assess potential selection bias related to unavailable follow-up fall information, baseline characteristics were compared between participants with and without classifiable fall status at the 3-year follow-up using standardized mean differences.

As a sensitivity analysis, inverse probability weighting was used to assess the robustness of the findings to differential availability of follow-up fall status. Separate weighting models were constructed for TUG, gait speed, and grip strength. For each analysis, the probability of having classifiable follow-up fall status was estimated from baseline age, sex, previous falls in the 12 months before baseline, comorbidity burden, cognitive status, opioid analgesic use, physical activity, frailty status, and the respective physical performance measure. Stabilized inverse probability weights were applied to re-estimate the association models, including the extended frailty-adjusted Model 3, which was used for comparison with the corresponding unweighted estimates.

Weight distributions, effective sample sizes, and covariate balance before and after weighting were examined to assess the performance of the weighting models. Primary regression and discrimination analyses were performed as complete-case analyses without imputation of missing values. For estimation of the inverse probability weights only, missing baseline covariate information was accommodated using missingness indicators; missing continuous covariate values were replaced by the sample median solely to permit estimation of follow-up probabilities. These replacements were not used in the outcome regression or discrimination analyses.

All statistical tests were two-sided, and $p < .05$ was considered statistically significant. All analyses were performed using R version 4.5.2 (R Foundation for Statistical Computing, Vienna, Austria).

Results

Study population

A total of 1,079 participants were included at baseline. Out of these, 822 participated in the 3-year follow-up. Among follow-up participants, 336 answered the trigger question regarding dizziness, balance problems, or falls affirmatively, 483 denied it, and 3 responded “don’t know”. Among participants with an affirmative trigger response, 90 reported one fall, 58 reported multiple falls, and 188 reported no falls in the detailed fall question. Thus, fall status for the 12 months preceding the 3-year follow-up interview was classifiable in 819 participants, including 671 non-fallers and 148 fallers (Fig. 1).

Baseline characteristics

Fallers were older and generally showed poorer physical performance and greater health impairment at baseline than non-fallers (Table 1). The largest standardized differences were observed for age, TUG, gait speed, frailty status, physical activity, and number of chronic conditions. Median TUG time was higher among fallers than non-fallers (10.9 vs. 9.4 s), whereas gait speed (1.0 vs. 1.1 m/s), grip strength (24.0 vs. 28.0 kg), and physical activity were lower. Previous falls at baseline were also more frequent among participants who reported falls at follow-up. Available sample sizes varied across baseline characteristics because of missing data.

Association analyses

In age- and sex-adjusted models (Model 1), longer TUG time and lower gait speed were associated with higher odds of reporting falls in the 12 months preceding the 3-year follow-up interview, whereas grip strength was not (Table 2). After additional adjustment for previous falls, comorbidity burden, cognitive status, opioid analgesic use, and physical activity (Model 2), the association remained for TUG and was attenuated for gait speed. Poorer TUG performance remained associated with reported falls across all three models, including after additional adjustment for frailty (Model 3). The association with slower gait speed was attenuated after further adjustment for frailty, whereas grip strength was not associated with reported falls in any model.

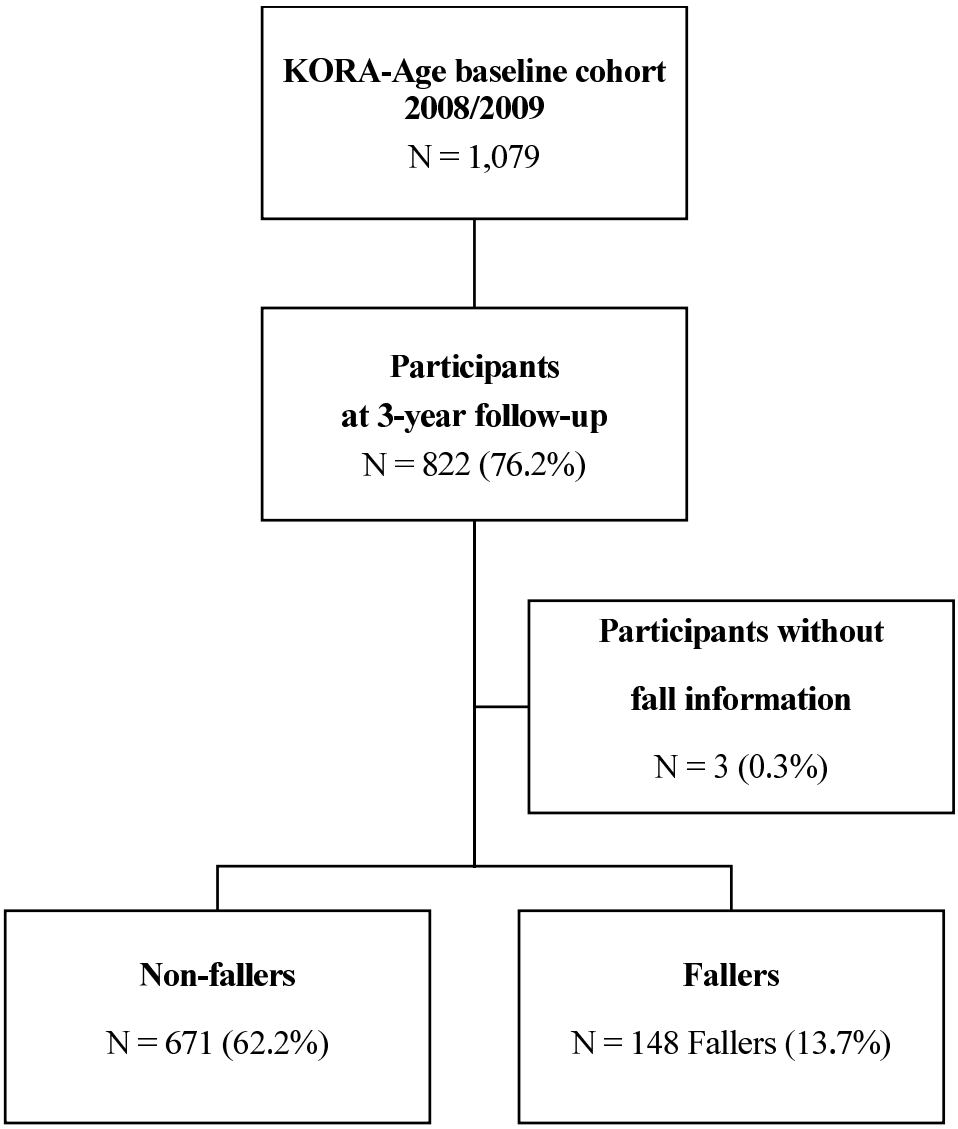

Fig. 1 Participant flow. Of 1,079 participants examined at baseline in 2008/2009, 822 participated in the 3-year follow-up. Fall status at follow-up for previous 12 months was classifiable for 819 participants, including 671 non-fallers and 148 fallers

Discriminative performance and incremental discrimination

In the common complete-case sample of 765 participants, including 135 fallers, the age- and sex-adjusted reference model had an AUC of 0.680 (95% CI 0.629–0.731). Addition of TUG resulted in an AUC of 0.700 (95% CI 0.650–0.750), gait speed in an AUC of 0.692 (95% CI 0.641–0.744), and grip strength in an AUC of 0.675 (95% CI 0.626–0.724). None of the physical performance measures significantly improved discrimination compared with age and sex alone (DeLong p=.067 for TUG, p=.205 for gait speed, and p=.598 for grip strength). The model including all three physical performance measures had an AUC of 0.700 (95% CI 0.650–0.751; DeLong p=.068 versus age and sex alone). Discrimination did not differ significantly between the TUG and gait-speed models (AUC 0.700 vs. 0.692; DeLong p=.356).

In a separate exploratory analysis, incremental discrimination beyond a clinical model including age, sex, previous falls, and opioid analgesic use was examined by sequentially adding TUG and the combined physical performance measures. Results and receiver operating characteristic curves are shown in Supplementary Tables 1 and Supplementary Fig. 1.

Table 1 Baseline characteristics by fall status

Characteristic	Available *N* overall	Overall	Non-fallers (*N* = 671)	Fallers (*N* = 148)	SMD
Age, years	819	75.0 (70.0–80.0)	74.0 (69.0–80.0)	79.0 (73.0–83.0)	0.613
Female sex, n (%)	819	406 (49.6)	318 (47.4)	88 (59.5)	0.244
Timed Up and Go, s	774	9.7 (8.5–11.6)	9.4 (8.4–11.1)	10.9 (9.5–14.3)	0.565
Gait speed, m/s	765	1.1 (0.9–1.3)	1.1 (1.0-1.3)	1.0 (0.8–1.1)	0.515
Grip strength, kg	815	28.0 (22.0–36.0)	28.0 (22.0–36.0)	24.0 (20.0–30.0)	0.386
Body mass index, kg/m²	819	27.9 (25.5–30.8)	27.9 (25.6–30.8)	28.1 (25.3–30.6)	0.017
Appendicular skeletal muscle mass index, kg/m²	779	7.2 (6.4–7.9)	7.2 (6.4–7.9)	6.9 (6.4–7.6)	0.199
Physical activity (PASE)	805	117.0 (82.0-156.0)	121.5 (88.0-161.0)	91.0 (62.0-132.0)	0.471
Vitamin D, ng/mL	805	20.3 (13.4–29.4)	20.7 (13.9–29.1)	18.4 (10.8–29.6)	0.151
Previous falls in the past 12 months *at baseline*, n (%)	818	126 (15.4)	86 (12.8)	40 (27.0)	0.361
Number of chronic conditions	815	2.0 (1.0–3.0)	2.0 (1.0–3.0)	3.0 (1.0–4.0)	0.442
Opioid analgesics use, n (%)	819	25 (3.1)	17 (2.5)	8 (5.4)	0.147
Frailty status	787		644	143	0.498
non-frail		490 (62.3)	429 (66.6)	61 (42.7)	
pre-frail		276 (35.1)	201 (31.2)	75 (52.4)	
frail		21 (2.7)	14 (2.2)	7 (4.9)	
EWGSOP2 sarcopenia status	815		668	147	0.225
No		715 (87.7)	595 (89.1)	120 (81.6)	
Probable		56 (6.9)	43 (6.4)	13 (8.8)	
Confirmed		44 (5.4)	30 (4.5)	14 (9.5)	
Cognitive status	793		653	140	0.110
normal		646 (81.5)	537 (82.2)	109 (77.9)	
MCI		109 (13.7)	86 (13.2)	23 (16.4)	
impaired		38 (4.8)	30 (4.6)	8 (5.7)	

Table 1: Baseline characteristics of fallers vs. non-fallers. Values are shown as median (interquartile range) for continuous variables and n (%) for categorical variables. For binary variables, the number and percentage of participants with the characteristic are shown. Percentages refer to participants with available data for the respective variable. Standardized mean differences (SMDs) compare fallers and non-fallers; for categorical variables with more than two categories, a single overall SMD is reported for the distribution across all categories. Abbreviations: EWGSOP2, European Working Group on Sarcopenia in Older People 2; MCI, mild cognitive impairment; PASE, Physical Activity Scale for the Elderly

Table 2 Associations of baseline physical performance measures with reported falls

Physical performance measure	Model	Adjusted OR (95% CI)	*p* value	*N*	Events
TUG, per 1-s increase	1	1.119 (1.059–1.182)	<0.001	774	140
	2	1.088 (1.024–1.157)	0.007	762	134
	3	1.095 (1.018–1.176)	0.014	761	134
Gait speed, per 0.1-m/s decrease	1	1.144 (1.047–1.251)	0.003	765	135
	2	1.102 (1.001–1.214)	0.049	754	130
	3	1.087 (0.979–1.206)	0.117	754	130
Grip strength, per 1-kg decrease	1	1.003 (0.972–1.035)	0.866	815	147
	2	0.999 (0.967–1.033)	0.968	790	139
	3	0.980 (0.945–1.017)	0.290	770	137

Table 2. Associations of baseline physical performance measures with reported falls. **Model 1** was adjusted for age and sex; **Model 2** additionally for previous falls at baseline, comorbidity burden, cognitive status, opioid analgesic use, and physical activity; and **Model 3** additionally for frailty status. ORs are expressed per worsening of physical performance: per 1-s increase in TUG time, per 0.1-m/s decrease in gait speed, and per 1-kg decrease in grip strength. Analyses were based on complete cases for the variables included in each model. CI, confidence interval; OR, odds ratio

Previously proposed TUG thresholds

In exploratory analyses of predefined TUG thresholds, participants above each threshold had higher age- and sex-adjusted odds of reporting falls at follow-up (Supplementary Table 2). For a threshold of > 12 s, sensitivity was 0.386 and specificity 0.825, with an age- and sex-adjusted OR of 1.92 (95% CI 1.24–2.96; $p = .003$). At higher thresholds, the proportion of participants classified above the threshold decreased, while specificity and positive predictive value increased and sensitivity decreased. Sensitivity was 0.271 for ≥ 13.5 s and 0.207 for > 15 s, whereas specificity increased to 0.897 and 0.935, respectively.

Attrition and sensitivity analyses

Fall status at follow-up could not be classified in 260 of the 1,079 baseline participants. Compared with participants with classifiable fall status, those without classifiable follow-up information were older and had poorer baseline physical performance, lower physical activity, greater comorbidity burden, and higher prevalence of frailty and sarcopenia (Supplementary Table 3). The largest standardized differences were observed for age and frailty status, indicating selective loss of follow-up information.

Inverse probability weighting achieved good covariate balance, with maximum absolute weighted SMDs of 0.042, 0.039, and 0.055 in the TUG, gait-speed, and grip-strength weighting models, respectively. Mean stabilized weights were approximately 1.0, and effective sample sizes ranged from 86.6% to 88.3% of the observed follow-up sample. Weighted estimates were similar to the primary analyses. In the frailty-adjusted models, the ORs were 1.09 (95% CI, 1.02–1.17) per 1-s increase in TUG, 1.10 (95% CI, 1.00-1.22) per 0.1-m/s decrease in gait speed, and 0.98 (95% CI, 0.95–1.02) per 1-kg decrease in grip strength. These findings suggest that accounting for measured predictors of follow-up availability had little impact on the observed associations, although residual selection bias related to unmeasured characteristics cannot be excluded.

Recurrent fallers

Among the 148 fallers, 58 participants reported recurrent falls at follow-up. In exploratory analyses using recurrent falls as the outcome, gait speed showed the highest single-marker AUC in the common complete-case subset (0.787), followed by TUG (0.776), whereas grip strength (0.739) did not improve discrimination beyond age and sex (Supplementary Table 4).

Discussion

Main findings

In this population-based cohort of community-dwelling older adults, poorer TUG performance remained associated with reported falls after adjustment for demographic and clinical characteristics, including frailty, whereas the association for gait speed was attenuated after additional adjustment for frailty and grip strength was not associated with falls. Despite these associations, the discriminatory performance of all three measures was modest, and none clearly improved discrimination beyond age and sex. Combining TUG, gait speed, and grip strength did not provide additional discriminatory value over TUG alone. Exploratory analyses of established TUG thresholds further showed a trade-off between increasing specificity and declining sensitivity at higher cut-offs.

Our findings are broadly consistent with previous studies showing associations between mobility-based performance measures and falls, while the discriminatory ability of individual physical performance tests has generally been limited. Kang et al. compared TUG, a 4-m walking test, and grip strength in community-dwelling older adults and reported only modest age- and sex-adjusted discrimination for *any* falls, whereas TUG performed

better for *recurrent* falls [7]. A later study from the same research group found that combining TUG, walking speed, and grip strength improved discrimination for recurrent falls but not for any falls [8].

In the KORA-Age cohort, Arvandi et al. previously reported an inverse association between grip strength and falls, but their analysis focused on grip strength and the potential mediating role of balance problems [5]. The present analysis extends previous work in KORA-Age by evaluating all three performance measures within the same population-based cohort and complements earlier comparative studies by formally examining incremental discrimination beyond a common reference model, combined test performance, and established clinical TUG thresholds.

Association versus discrimination

The distinction between association and discrimination is important for the interpretation of physical performance measures as potential fall-risk markers. TUG remained independently associated with reported falls after multivariable adjustment and after additional adjustment for frailty, but its addition to an age- and sex-based reference model resulted in only a small increase in AUC that did not reach statistical significance. The attenuation of the association with gait speed after additional adjustment for frailty should be interpreted cautiously. The Fried frailty phenotype incorporates slow walking speed, low grip strength, and low physical activity; adjustment for frailty may therefore partly account for functional characteristics already captured by gait speed and grip strength themselves. A similar overlap applies to grip strength, which showed neither an independent association nor incremental discriminatory value in the present analyses.

The descriptive difference in grip strength between fallers and non-fallers was largely attenuated after adjustment for age and sex, suggesting that part of this difference reflected differences in the demographic composition of the groups rather than an independent association with falls. Therefore, grip strength may capture broader aspects of muscle function and general vulnerability rather than mobility-specific characteristics directly related to falls. Moreover, the discriminatory performance of TUG did not differ significantly from that of gait speed, and combining all three performance measures did not improve discrimination beyond TUG alone.

These findings suggest that physical performance measures reflect functional characteristics associated with falls but have limited ability, either individually or in combination, to distinguish between older adults with and without falls. This interpretation is consistent with systematic reviews showing limited stand-alone utility of performance-based measures for identifying older adults who fall [24, 25]. Exploratory analyses showed numerically higher AUCs for recurrent falls, particularly for gait speed and TUG, although the small number of recurrent fallers limits firm conclusions.

Clinical interpretation of TUG thresholds

The exploratory threshold analyses further illustrate the limitations of using TUG as a stand-alone screening tool for falls. At the > 12-s threshold, specificity was substantially higher than sensitivity, and increasing the cut-off to 13.5 and 15 s further increased specificity while markedly reducing sensitivity. Thus, higher thresholds identified smaller groups with a higher prevalence of falls but missed an increasing proportion of individuals who reported falls. This pattern is consistent with the meta-analysis by Barry et al., which likewise found substantially lower sensitivity than specificity at the ≥ 13.5-s threshold, as well as with findings by Kang et al. showing increasing specificity at more restrictive TUG cut-offs [7, 26]. Rather than supporting a new ‘optimal’ threshold, our findings suggest that established TUG cut-offs may be useful for identifying functionally vulnerable subgroups but are insufficient as a sensitive screening strategy for falls.

Strengths and limitations

The study's strengths lie in its population-based cohort, the standardization of the assessment of three widely utilized physical performance measures, and the evaluation of these within a unified analytical framework. This framework addresses associations, incremental discrimination, combined performance, and predefined clinical thresholds.

It is imperative to acknowledge the limitations of the present study. Falls were self-reported for the 12 months preceding the follow-up interview rather than prospectively recorded throughout the 3-year interval, introducing the possibility of recall bias and leaving falls occurring earlier in the interval unascertained. Furthermore, alterations in mobility, health status, or other fall-related factors between the baseline assessment and the recall period may have influenced the observed associations. Fall ascertainment followed a two-step questionnaire structure in which participants who denied the initial screening question were not administered the detailed fall question. Therefore, some participants with falls may have been misclassified as non-fallers. Follow-up fall status was unavailable for 260 baseline participants, who were older and more functionally vulnerable, indicating selective attrition. However, inverse-probability weighting yielded estimates very similar to the primary analyses, suggesting that accounting for measured predictors of follow-up availability had little impact on the observed associations, although residual selection bias cannot be excluded. Despite adjustment for multiple relevant covariates, residual and unmeasured confounding remains possible. Finally, proxy interviews were permitted when participants were unable to complete the interview themselves. A published description of the KORA-Age Age2 follow-up reported that 29 of 822 interviews (3.5%) were conducted with proxies [19]. Because participant-level information on proxy status was not available in our analytical dataset, we could not determine the number of proxy interviews among participants included in the present analysis or assess whether fall reporting differed between self- and proxy-reported interviews.

Conclusions

In this population-based cohort of older adults, TUG remained associated with self-reported falls in the 12 months preceding the 3-year follow-up interview after multivariable and additional frailty adjustment. However, none of the three performance measures showed statistically significant incremental discrimination beyond age and sex, and combining the measures did not improve discrimination over TUG alone. Higher established TUG thresholds identified smaller groups with a higher prevalence of reported falls but progressively lower sensitivity. These findings do not support the use of TUG alone as a sensitive stand-alone screening tool for falls.

Abbreviations

ASMM	appendicular skeletal muscle mass
AUC	area under the curve
BIA	bioelectrical impedance analysis
BMI	body mass index
CI	confidence interval
NICE	National Institute for Health and Care
OR	odds ratio
PASE	Physical Activity Scale for the Elderly
ROC	receiver operating characteristic
SMD	standardized mean difference
TICS	m-modified Telephone Interview for Cognitive Status
WHO	World Health Organization

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1186/s12877-026-08627-9.

Acknowledgements Our research group thanks all participants for their long-term commitment to the KORA study, the staff for data collection and research data management and the members of the KORA Study Group (https://www.helmholtz-munich.de/en/epi/cohort/kora) who are responsible for the design and conduct of the study.

Author contributions Conceptualization: BT, MD, MTH, SMData curation: BT, MH, MTHInvestigation: CM, MD, MTH, SM, UFMethodology: BT, CM, MD, MH, MTH, SM, UFSupervision: MD, RSWriting-original draft: CM, MD, RS, SMWriting-review and editing: all.

Funding The KORA study was initiated and financed by the Helmholtz Zentrum München - German Research Center for Environmental Health, which is funded by the German Federal Ministry of Research, Technology and Space (BMFTR) and by the State of Bavaria. Data collection in the KORA study is done in cooperation with the University Hospital of Augsburg. The KORA-Age project was financed by the German Federal Ministry of Education and Research (BMBF FKZ 01ET0713 and 01ET1003A) as part of the 'Health in old age' program.

Data availability KORA data and biosamples are available upon request by means of a project agreement subject to approval by the KORA Board. Via the KORA.PASST tool, scientists can view and download data dictionaries of the different KORA surveys and submit a proposal. Please apply for KORA data or biosamples at the KORA.PASST website: https://helmholtz-muenchen.managed-otrs.com/external/ .

Declarations

Ethics approval and consent to participate All participants provided written informed consent prior to inclusion, and all procedures were conducted in accordance with the Declaration of Helsinki. The study was approved by the Ethics Committee of the Bavarian Medical Association (reference no. 08064).

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Chen Y, Dai F, Huang S, Qi D, Peng C, Zhang A, et al. Global, regional, and national burden of falls among older adults: findings from the Global Burden of Disease Study 2021 and Projections to 2040. NPJ Aging. 2025;11(1):85. https://doi.org/10.1038/s41514-025-00275-4.
2. Jochum D, Siegel SD, Ploschberger G, Fenske L, Warneke K. Resistance training effects on gait parameters in older adults-a systematic review with multilevel meta-analysis and meta-regression. Geroscience. 2026. https://doi.org/10.1007/s11357-026-02156-4.
3. Neilson J, MacIntyre D, Whitney J, Aung T, Lally M, Guideline C, et al. Assessment and prevention of falls in older people and in people 50 and over at higher risk-summary of updated NICE guidance. BMJ. 2026;392:s223. https://doi.org/10.1136/bmj.s223.
4. Organization WH. WHO global report on falls prevention in older age. 2007:1–53.
5. Arvandi M, Strasser B, Volaklis K, Ladwig KH, Grill E, Matteucci Gothe R, et al. Mediator Effect of Balance Problems on Association Between Grip Strength and Falls in Older Adults: Results From the KORA-Age Study. Gerontol Geriatr Med. 2018;4:2333721418760122. https://doi.org/10.1177/2333721418760122.

6. Viccaro LJ, Perera S, Studenski SA. Is timed up and go better than gait speed in predicting health, function, and falls in older adults? J Am Geriatr Soc. 2011;59(5):887–92. https://doi.org/10.1111/j.1532-5415.2011.03336.x.
7. Kang L, Han P, Wang J, Ma Y, Jia L, Fu L, et al. Timed Up and Go Test can predict recurrent falls: a longitudinal study of the community-dwelling elderly in China. Clin Interv Aging. 2017;12:2009–16. https://doi.org/10.2147/CIA.S138287.
8. Wang L, Song P, Cheng C, Han P, Fu L, Chen X, et al. The Added Value of Combined Timed Up and Go Test, Walking Speed, and Grip Strength on Predicting Recurrent Falls in Chinese Community-dwelling Elderly. Clin Interv Aging. 2021;16:1801–12. https://doi.org/10.2147/CIA.S325930.
9. Linkohr B, Heier M, Gieger C, Thorand B, Grallert H, Holle R, et al. Cohort Profile: Cooperative Health Research in the Region of Augsburg (KORA) 1984–2024. Int J Epidemiol. 2025;54(6). https://doi.org/10.1093/ije/dyaf187.
10. Peters A, Doring A, Ladwig KH, Meisinger C, Linkohr B, Autenrieth C, et al. [Multimorbidity and successful aging: the population-based KORA-Age study]. Z Gerontol Geriatr. 2011;44(Suppl 2):41–54. https://doi.org/10.1007/s00391-011-0245-7.
11. Washburn RA, Smith KW, Jette AM, Janney CA. The Physical Activity Scale for the Elderly (PASE): development and evaluation. J Clin Epidemiol. 1993;46(2):153–62. https://doi.org/10.1016/0895-4356(93)90053-4.
12. Kirchberger I, Meisinger C, Heier M, Zimmermann AK, Thorand B, Autenrieth CS, et al. Patterns of multimorbidity in the aged population. Results from the KORA-Age study. PLoS ONE. 2012;7(1):e30556. https://doi.org/10.1371/journal.pone.0030556.
13. Conzade R, Grill E, Bischoff-Ferrari HA, Ferrari U, Horsch A, Koenig W, et al. Vitamin D in Relation to Incident Sarcopenia and Changes in Muscle Parameters Among Older Adults: The KORA-Age Study. Calcif Tissue Int. 2019;105(2):173–82. https://doi.org/10.1007/s00223-019-00558-5.
14. Autenrieth CS, Karrasch S, Heier M, Gorzelniak L, Ladwig KH, Peters A, et al. Decline in gait performance detected by an electronic walkway system in 907 older adults of the population-based KORA-Age study. Gerontology. 2013;59(2):165–73. https://doi.org/10.1159/000342206.
15. Sergi G, De Rui M, Veronese N, Bolzetta F, Berton L, Carraro S, et al. Assessing appendicular skeletal muscle mass with bioelectrical impedance analysis in free-living Caucasian older adults. Clin Nutr. 2015;34(4):667–73. https://doi.org/10.1016/j.clnu.2014.07.010.
16. Cruz-Jentoft AJ, Bahat G, Bauer J, Boirie Y, Bruyere O, Cederholm T, et al. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing. 2019;48(4):601. https://doi.org/10.1093/ageing/afz046.
17. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol Biol Sci Med Sci. 2001;56(3):M146–56. https://doi.org/10.1093/gerona/56.3.m146.
18. Cook SE, Marsiske M, McCoy KJ. The use of the Modified Telephone Interview for Cognitive Status (TICS-M) in the detection of amnestic mild cognitive impairment. J Geriatr Psychiatry Neurol. 2009;22(2):103–9. https://doi.org/10.1177/0891988708328214.
19. Steinbeisser K, Schwarzkopf L, Schwettmann L, Laxy M, Grill E, Rester C, et al. Association of physical activity with utilization of long-term care in community-dwelling older adults in Germany: results from the population-based KORA-Age observational study. Int J Behav Nutr Phys Act. 2022;19(1):102. https://doi.org/10.1186/s12966-022-01322-z.
20. Prevention CfDCa. National Center for Health Statistics. National Health and Nutrition Examination Survey 2003–2004: Balance (BAQ_C). 2005.
21. Larsson BAM, Johansson L, Johansson H, Axelsson KF, Harvey N, Vandenput L, et al. The timed up and go test predicts fracture risk in older women independently of clinical risk factors and bone mineral density. Osteoporos Int. 2021;32(1):75–84. https://doi.org/10.1007/s00198-020-05681-w.
22. Montero-Odasso M, van der Velde N, Martin FC, Petrovic M, Tan MP, Ryg J, et al. World guidelines for falls prevention and management for older adults: a global initiative. Age Ageing. 2022;51(9). https://doi.org/10.1093/ageing/afac205.
23. Shumway-Cook A, Brauer S, Woollacott M. Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test. Phys Ther. 2000;80(9):896–903.
24. Lusardi MM, Fritz S, Middleton A, Allison L, Wingood M, Phillips E, et al. Determining Risk of Falls in Community Dwelling Older Adults: A Systematic Review and Meta-analysis Using Posttest Probability. J Geriatr Phys Ther. 2017;40(1):1–36. https://doi.org/10.1519/JPT.0000000000000099.
25. Schoene D, Wu SM, Mikolaizak AS, Menant JC, Smith ST, Delbaere K, et al. Discriminative ability and predictive validity of the timed up and go test in identifying older people who fall: systematic review and meta-analysis. J Am Geriatr Soc. 2013;61(2):202–8. https://doi.org/10.1111/jgs.12106.
26. Barry E, Galvin R, Keogh C, Horgan F, Fahey T. Is the Timed Up and Go test a useful predictor of risk of falls in community dwelling older adults: a systematic review and meta-analysis. BMC Geriatr. 2014;14:14. https://doi.org/10.1186/1471-2318-14-14.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Sebastian Martini[1] · Michael Drey[1] · Ralf Schmidmaier[1] · Uta Ferrari[1] · Canan Marcon[1] · Annette Peters[2,3] · Marie-Theres Huemer[2] · Margit Heier[2,4] · Barbara Thorand[2,4]

✉ Sebastian Martini
sebastian.martini@med.uni-muenchen.de

1 Department of Medicine IV, Geriatrics, LMU Munich Hospital, Ludwig-Maximilians-Universität München, Munich, Germany

2 Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health (GmbH), Neuherberg, Germany

3 Institute for Medical Information Processing, Biometry and Epidemiology (IBE), Faculty of Medicine, Ludwig-Maximilians-Universität München, Pettenkofer School of Public Health, Munich, Germany

4 KORA Study Centre, University Hospital of Augsburg, Augsburg, Germany