

Validity and Inter-observer Reliability of the TURN 180 Test to Identify Older Adults Who Reported Falls

Yulia Gamerman MPT¹, Moshe Hoshen MD², Avner Herman Cohen MD^{4,5}, Zhana Alter PT³, Luzit Hadad PT³ and Itshak Melzer PT PhD¹

¹Department of Physical Therapy, Recanati School for Community Health Professions, Faculty of Health Sciences, Ben-Gurion University of the Negev, Beer Sheva, Israel

²Clalit Research Institute, Chief Physician's Office, Clalit Health Services, Tel Aviv, Israel

³Department of Physiotherapy, Clalit Health Services, Petah Tikva, Israel

⁴Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

⁵Pediatric Ambulatory Community Clinic, Petah Tikva, Israel

ABSTRACT: **Background:** Falls while turning are associated with increased risk of hip fracture in older adults. Reliable and clinically valid methods for turn ability assessments are needed.

Objectives: To explore the inter-observer reliability and known group validity of the TURN 180 test.

Methods: We divided 78 independent older adults (mean age 76.6 ± 6.5 years) into three groups: non-fallers, infrequent fallers (1–2 falls per year), and recurrent fallers (> 2 falls per year). Participants underwent performance-based tests: Timed Up and Go (TUG), Performance Oriented Mobility Assessment (POMA), and Berg Balance Scale (BBS). TUG was videotaped for later analysis of the TURN 180 test by two blinded observers.

Results: A significant difference was found in the TURN 180 test parameters among the groups ($P < 0.04$). TURN 180 was highly correlated with TUG ($r = 0.81–0.89$, $P < 0.001$) and BBS ($r = -0.704–0.754$, $P < 0.0001$) and moderately with POMA ($r = -0.641–0.698$, $P < 0.0001$). The number of steps was found to be the strongest parameter to determine fallers among older adults (specificity 96.3%, sensitivity 40%). Inter-rater reliability (intraclass correlation coefficient 0.91–0.96, $P < 0.0001$) was found to be excellent for the number of steps, time taken to accomplish a turn, and total test score categories.

Conclusions: The TURN 180 test is highly reliable and can identify the older adults who fall. Our results show that the TURN 180 test can serve as a good performance-based examination for research or clinical setting.

KEY WORDS: aging, balance, fall risk, TURN 180 test

IMAJ 2019; 21: 269–274

fracture. It is estimated that 95% of hip fractures are the result of falls, 40% of hip fractures occur while walking, and 18% occur during turning [1]. A fall while turning is 7.9 times more likely to result in a hip fracture than a fall while walking straight [3]. This finding suggests that turning is a greater challenge for older adults than walking straight and may carry a significant risk for hip fracture.

Research on turning has primarily been limited to laboratory or clinical settings, although many daily activities require turning. Glaister and co-workers [4] found that non-straight steps during walking comprise approximately 35 to 45% of all steps taken in a typical day. Turn performance testing of older adults while walking has been used to distinguish fallers from non-fallers. Dite et al. [5] found that older adults who had recurrent falls during the previous 6 months had to take two or more steps to complete a 180-degree turn compared to older adults who reported a single fall. Specific turn characteristics (i.e., turn time, turn steps, total turn score) were found to discriminate between groups of healthy and impaired older adults and had sufficient sensitivity to identify older adults who reported recurrent falls. Thigpen and colleagues [6] found that the older adults who had difficulty turning took more steps, needed more time to complete the turn, and usually did not use a pivot type of turn compared to young adults and older adults with no difficulty in turning.

The TURN 180 test is a quick, practical, and low-cost test that may be able to identify fallers and may estimate a change following an intervention. The TURN 180 test can be assessed during the more common Timed up and Go test (TUG), and it takes only a few more minutes to observe and analyze a video clip. The TURN 180 test assesses an important component of a function in daily life that has, to date, not been sufficiently addressed in research.

The primary purpose of the current study was to evaluate the inter-observer reliability of the TURN 180 test. Our second aim was to explore the ability of the TURN 180 test to distinguish between older adults who were classified as

As the geriatric population continues to grow, the consequences of falls and fall-related injuries increases (e.g., hip fracture, wrist fractures, and head injuries) [1]. Aizen and colleagues [2] reported that recurrent falls before a hip fracture-related fall is associated with substantial loss of functional independence and adversely affects rehabilitation outcome of hip

non-fallers, infrequent fallers, or recurrent fallers. Our third aim was to examine the associations between gold standard performance-based tests of balance and the TURN 180 test. We suggested that the inter-observer reliability of TURN 180 test parameters would be high, especially for number of steps and time to complete the test, and somewhat lower for turn strategy and turn stability. We proposed that the number of steps and time to complete the TURN 180 test in recurrent fallers would be higher than for infrequent fallers and non-fallers. In addition, the TURN 180 test would demonstrate high associations with the TUG, Berg Balance Scale test (BBS), and Performance Oriented Mobility Assessment (POMA) because those tests also examine a person's ability to turn.

PATIENTS AND METHODS

A convenience sample of 78 older adults (13 men and 65 women, mean age 76.62 ± 6.5 years, range 65–90 years) was recruited. Patients were included if they were 65 years of age or older and able to ambulate independently for a distance of at least 10 meters. The exclusion criteria were visual impairment that could not be corrected by glasses, acute cardiopulmonary conditions (myocardial infarction, acute asthma) within the previous 3 months, severe congestive heart failure, neurologic pathologies (stroke, recurrent sclerosis, Parkinson's disease), amputation of lower limb, vertigo, acute low back pain, acute ankle/knee/hip joint inflammation, or arthritis.

The participants were divided into three groups according to the number of self-reported falls in the previous 12 months: 27 reported no falls, 26 reported one to two falls, and 25 reported more than two falls. After the participant signed informed consent documents, experienced physical therapists conducted the assessments and collected the demographical data, information on the general health condition, fall history, and level of physical activity (measured in hours per week) through the short version International Physical Activity Questionnaires (IPAQ) [7]. The short IPAQ comprises a set of questions that provide information about the kinds of physical activities that people do as part of their everyday lives and the time they spent being physically active in the last 7 days (i.e., working, doing house and yard work, getting from place to place, and participating in recreation, exercise, or sport). In addition, the patients underwent three performance-based balance tests: TUG, POMA, and BBS.

To observe the subjects turning abilities while walking, we used a standardized study protocol. Each participant was given one practice trial and then performed the TUG three successive times while being videotaped. To establish the reliability of the TURN 180 test categories, the TUG video clips of the testing procedure were later evaluated separately by two physical therapists who were blinded to each other's assessment. They analyzed the TURN 180 test during the TUG video clips to

determine the turn strategy, stability, number of steps taken during the turn, and the time needed to accomplish the turn using the slow-motion and stop-action capabilities of the video system. The parameters that we evaluated were gradation of staggering, number of steps, time to accomplish turn and turn strategy according to Thigpen et al. [6]. The mean value of three trials for each rater was considered. A 180-degree turn during the TUG is defined as the beginning and end of the 180 reversal of direction at the turn line on the floor while walking. The last heel-strike prior to initiation of the reversal of direction is designated as the beginning of the turn. The heel-strike of the first step progressing in a direct line back to the chair is designated the end of the turn.

The study was approved by the ethics committee of Clalit Health Services (0003-15-COM).

SAMPLE SIZE ESTIMATION

Sample size requirements were calculated based on the number of steps needed to complete the Turn 180 test. We also used the time to accomplish the turn in older adults who had no difficulty in turning and older adults who had difficulty in turning [6]. For both calculations, the probability of a type I error was 0.05, and probability of a type II error was 0.20. Using net differences values for the groups in the number of steps needed to complete the Turn 180 test and the time to accomplish the turn in combination with the initial variance estimates, it was determined that 19 participants per group would be required.

DATA AND STATISTICAL ANALYSIS

Statistical analyses were performed using Statistical Package for the Social Sciences software version 16 (SPSS Inc., Chicago, IL, USA). A descriptive analysis of the groups' characteristics for nominal variables is presented by frequencies or percentage and is compared within the groups by chi-square or the Fisher's exact test. Continuous variables are presented by the mean values and standard deviations and ordinal variables are presented by the median and inter-quartile range (IQR).

To differentiate inter-observer variation, a two-way random model for inter-class correlation coefficients (ICC_2) [8] was used assuming that each subject was assessed by two different observers, in random order, that reliability was calculated from an average measurement. Reliability is considered high with $ICC > 0.80$, substantial for $ICCs 0.61-0.80$, moderate for $ICCs 0.41-0.60$, and poor to fair for $ICCs < 0.4$ [9].

In addition, mountain plot graphs for the TURN 180 parameters were generated separately for staggering, number of steps, number of steps grade, turn time (seconds), turn time grade, strategy (seconds), and total score [Figure 1]. The mountain plot graphs display the cumulative distribution of the difference between the scores in the two testing procedures. The graph was created by computing a percentile rank for ranked difference between the respondent scores and folded at the 50th percentile

rank (percentile rank is defined as the proportion of cases having lower or equal value to the score under consideration). The mountain graph plots cumulative percentages (y-axis values) against the ranked difference scores (x-axis values). In folding the graph, the percentile ranks for different scores above the 50th percentile were obtained by subtracting the actual percentile rank from 100 [10]. The mountain plot has several advantages, including locating the median immediately, easily determining symmetry, and observing outliers to determine central or tail percentiles. These advantages are especially useful in examining scoring differences for each TURN 180 parameter scored by the first and second testing procedure.

To test the ability of the TURN 180 test to identify non-fallers, infrequent fallers, and recurrent fallers, a one-way ANOVA was performed or Kruskal–Wallis test if the data were not normally distributed. If there were significant differences, additional post-analysis least significant difference was performed or the Mann–Whitney test was used in case the data were not normally distributed. Statistical significance was set at $P < 0.05$.

The Spearman correlation was performed to determine the associations between TURN 180 test parameters and TUG, POMA, and BBS scores. Correlation magnitude was estimated as absent to little ($r = 0.00–0.25$), low ($0.26–0.49$), moderate ($0.50–0.69$), high ($0.70–0.89$), or very high ($0.90–1.00$). In addition, a logistic regression model adjusted for age was built to find the strongest variable to determine fallers (recurrent or infrequent fallers) from non-fallers. We compared the goodness of fit of the logistic models by using the likelihood-ratio test and receiver operating characteristics (ROC) analysis in the tests.

RESULTS

There were no significant differences in the groups with regard to body mass index and specific disease existence. The recurrent fallers were significantly older (mean 80.3 ± 5.9 years) than non-fallers (75.0 ± 6.2 years, $P = 0.007$) or infrequent fallers (74.8 ± 6.2 , $P = 0.006$) and had significantly more comorbidities ($P = 0.001$) than the two other groups. The recurrent fallers were taking significantly more medications than infrequent fallers ($P = 0.003$) or non-fallers ($P < 0.001$). There was a significant difference in gender proportions in the groups ($P = 0.03$). In the non-fallers, 29.6% were men compared to 3.8% in infrequent fallers and 16% in recurrent fallers. The recurrent fallers had a significantly lower level of physical activity than the non-fallers (mean = 1.34 hours/week vs. 2.72 hours/week, $P = 0.036$). Of the recurrent fallers, 40% used a walking stick compared with 7.7% of the non-fallers and 7.4% of the infrequent ($P = 0.0008$) fallers.

High inter-observer agreement was found for the number of steps (quantity) and time taken to accomplish a turn (in seconds) and the total score categories (ICC 0.91–0.96, $P < 0.0001$). Moderate to substantial agreement was found for

strategy, number of steps taken, and time taken to accomplish a turn measured by grades (ICC 0.5–0.8, $P < 0.0001$). Slight to fair agreement was found for the staggering in the TURN 180 categories.

To gain an appreciation of the direction and magnitude of the score differences between two observers' testing procedures, we used mountain plots for TURN 180 test parameters. Because the differences in scores were computed using the observer A value scores minus the observer B scores, a positive difference in scores could be interpreted such that the observer A estimates the performance at a higher score than observer B and vice versa. For staggering, the median difference score was near zero, for the number of steps the estimate was near -0.7, for the number of steps grade it was near -0.75–0, for turn time it was near zero, for strategy it was near zero and for total score it was near -0.25. The median difference scores for all TURN 180 test parameters were near zero. These results indicate that no large systematic differences were detected between two observers [Figure 1].

The average values of the four categories (staggering, number of steps, time taken to accomplish the turn, strategy) and total score of the TURN 180 test of the three groups were compared. A significant difference was found between the groups in all categories ($P < 0.001$) except staggering ($P = 0.302$) [Table 1].

There were also significant differences for the groups in BBS ($P < 0.001$), POMA ($P < 0.001$), and TUG ($P < 0.0001$). High correlation was found in TURN 180 test parameters (number of steps, time taken to accomplish the turn, strategy), the BBS total score ($r = -0.704–0.754$, $P < 0.0001$), and TUG ($r = 0.81–0.89$, $P < 0.001$). A moderate correlation was found with the POMA ($r = -0.641–0.698$, $P < 0.0001$). Staggering category showed a poorer yet significant correlation with the BBS, TUG, and POMA [Table 2].

Forward stepwise logistic regression was used to identify the strongest variables that can determine recurrent fallers or infrequent fallers from non-fallers. We found that the model that included the TURN 180 category number of steps grade, OR [odds ratio] = 16.56, $P < 0.0001$, 95% confidence inter-

Table 1. Mean values \pm standard deviation of the TURN 180 test categories among the groups (non-fallers, infrequent fallers, and recurrent fallers)

Category	Non-fallers	Infrequent fallers	Recurrent fallers	P value
Staggering	0.1 \pm 0.3	0.2 \pm 0.4	0.4 \pm 0.8	0.302
Number of steps	3.2 \pm 1.3	4.2 \pm 1.4	5.4 \pm 1.6	< 0.0001
Number of steps, grade	0.7 \pm 0.5	1.2 \pm 0.5	1.7 \pm 0.4	< 0.0001
Time to accomplish turn, seconds	2.3 \pm 0.7	2.8 \pm 1	3.8 \pm 1.5	< 0.0001
Time taken to accomplish turn, grade	0.5 \pm 0.6	0.9 \pm 0.8	1.5 \pm 0.7	< 0.0001
Strategy	0.7 \pm 0.5	1.1 \pm 0.5	1.5 \pm 0.5	< 0.0001
Total score	2.0 \pm 1.5	3.4 \pm 1.9	5.1 \pm 2.0	< 0.0001

Figure 1. Mountain plots for agreement between two observers, mean scores from three trials: staggering, number of steps, number of steps grade, turn time (seconds), turn time grade, strategy (seconds), and total score

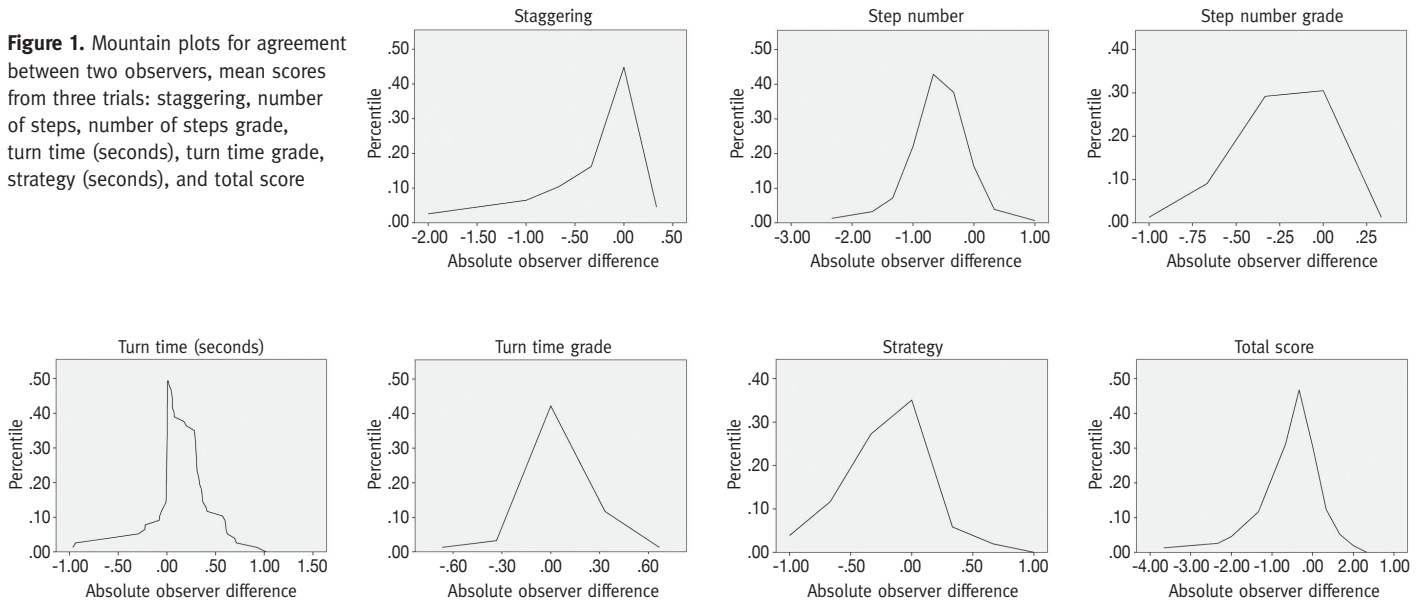


Table 2. Spearman correlation for TURN 180 test categories, mean values, and TUG, POMA, and BBS tests

Category of TURN 180 test	TUG	POMA	BBS
Staggering	0.377	-0.384	-0.292
Number of steps	0.833	-0.695	-0.735
Number of steps, grade	0.835	-0.671	-0.727
Time taken to accomplish turn, seconds	0.893	-0.698	-0.731
Time taken to accomplish turn, grade	0.864	-0.668	-0.704
Strategy	0.810	-0.641	-0.712
Total score	0.881	-0.697	-0.754

P < 0.001 for all correlations

BBS = Berg Balance Scale test, POMA = Performance Oriented Mobility Assessment, TUG = Timed get Up and Go test

Table 3. Comparison of goodness of fit of the regression models for number of steps grade and activity level, TUG, POMA, and BBS tests, as predictors of falling

Model variables	Odds ratio (95% confidence interval)	<i>P</i> value	Area under the curve
Step number grade	16.56 (3.84–71.45)	< 0.0001	0.849
Physical activity level (IPAQ score)	0.75 (0.57–0.99)	0.04	
Step number grade	15.89 (3.78–62.57)	< 0.0001	0.809
TUG	1.4 (1.12–1.75)	0.003	0.813
BBS	0.78 (0.67–0.91)	0.002	0.77
POMA	0.72 (0.54–0.96)	0.03	0.679

BBS = Berg Balance Scale test, IPAQ = International Physical Activity Questionnaires, POMA = Performance Oriented Mobility Assessment, ROC = receiver operating characteristics, TUG = Timed get Up and Go test

val [95%CI] (3.84–71.45) and level of physical activity, OR = 0.75, *P* = 0.041, 95%CI (0.57–0.99), adjusted for age, effectively detected the fallers [Table 3].

To determine the cutoff point of number of steps to distinguish fallers (at least one fall) from non-fallers, we successively divided the study population according to the number of steps taken to accomplish the turn and evaluated the presence of fallers in each group. We found that, at step 5, the rate of fallers rises to almost 100% while only 60% at step 4. To compute the specificity and sensitivity of the cutoff point (above four steps), we separated our population into two groups according to the cutoff point: ≤ 4 steps and ≥ 5 steps. We performed a cross-tab analysis with a faller/non-faller variable. The specificity rate at the cutoff point was 96.3% and sensitivity was 40%.

DISCUSSION

In the present study, we found excellent ICCs for the number of steps, time taken to accomplish the turn, and total score test categories. These findings are consistent with Thigpen's group [6]. They found substantial agreement on the number of steps taken during the turn and an almost perfect agreement for time taken to accomplish the turn with regard to quantitative items and mountain plots. Our study adds to existing knowledge by showing that there are no systematic differences between two observers. Median difference scores for all TURN 180 test parameters were near zero, indicating high stability and no large systematic differences between the observers. The staggering and strategy, however, are parameters that are related to the observer's subjective evaluation (i.e., qualitative items) and were found to be less reliable.

The excellent inter-rater agreement of the TURN 180 test found in the present study (ICC 0.91–0.96) is similar to the BBS (ICC > 0.91) [11] and TUG (ICC 0.56–0.99) [12,13] and somewhat higher than reported for POMA (ICC 0.72–0.86) [14]. This finding suggests that older adults who have difficulty in turning tend to fall more often. It is consistent with a previous study [5] that showed that turn step number and time taken to turn parameters have good sensitivity in identifying recurrent fallers only but not infrequent fallers.

To the best of our knowledge, no research has been undertaken to study the correlations among TURN 180, TUG, BBS, and POMA, which are three well-established balance performance tests. Our results show high correlation with TUG and BBS, and a moderate correlation with POMA. These outcomes are probably due to fact that TUG is part of the TURN 180 test, while part of the POMA is dedicated to the gait symmetry and balance, which is not assessed by the TUG, BBS, and TURN 180 tests. Dite's team [5] found a moderate to high correlation ($r = -0.62$ – -0.76) between the quantitative items used in the TURN 180 and BBS tests.

In the two-parameter model, the number of steps and activities were able to best determine the fallers. Individuals with a low level of physical activity and a higher grade in the number of steps while turning, respectively, had a greater chance of being a faller. The results showed that taking more than four steps for a turn identified fallers (both recurrent and infrequent) and non-fallers with excellent specificity (96.3%) and low sensitivity (40%). Accordingly, if an older adult is taking four steps or fewer to turn, we can conclude with 96% confidence that he/she is not a faller. However, Dite et al. [5] showed a higher level of sensitivity for the turn step number (81–92%) in identifying recurrent fallers. This result could be due to the different definitions of the reported number and period of falls in each group. Nevitt and colleagues [15] showed that older adults who take five or more steps to complete the turn are almost twice as likely to sustain two or more falls in the following year compared with people who take fewer than five steps. In view of these findings, we suggest that using only the number-of-steps parameter is sufficient to determine faller (recurrent or infrequent) versus non-faller status.

Comparing the specificity and sensitivity of the TURN 180 test to those of the other three tests conducted in this study, we found that the specificity of the TURN 180 test was similar to that of the BBS (82–96%) [16,17] and higher than the TUG (60%) [18] and POMA (78%) [14,19]. The sensitivity of the TUG (56%) [19], BBS (53–91%) [16,17], and POMA (68%) [20] tests is higher than the TURN 180 test, suggesting that the latter is as good as BBS and better than the two other tests in identifying older adults who are non-fallers. It is important to note that measuring the number of steps during a 180 degree turn takes several seconds, whereas BBS takes approximately 20 minutes. Consequently, the TURN 180 test is similarly effective,

but takes significantly less time. To determine fallers rather than non-fallers, the other three tests should be performed.

There were several limitations to this study. First, the distributions of the groups depended on the retrospective self-report of the participants about the last year. Second, there is a causality problem. Based on our current research, we do not know whether the decreased ability to turn can be the cause of the falls or being a faller decreases the confidence of people to move, and therefore, they tend to turn more slowly and add extra steps. The common characteristics of the recurrent fallers are older age, high number of co-morbidities, and medication in the current study. All of these factors might affect the performance of the TURN 180 test. Third, while compelling, the results are not completely conclusive because a larger sample needs to be assessed to draw more solid conclusions.

CONCLUSIONS

The TURN 180 test is a quick, practical, and low-cost test that can be used to identify non-fallers. The TURN 180 test parameters, time to accomplish a turn, the number of steps, turn strategy, and total score are reliable and valid assessments, which are able to distinguish between older adults with different numbers of falls. The staggering test parameter has low reliability and validity and thus should not be used as a parameter. The TURN 180 test has high correlation with TUG and BBS, but moderate correlation with POMA. The number of steps was found to be the strongest parameter to determine fallers among older adults, suggesting that a person who takes more than four steps during a turn tends to be a faller.

Correspondence

Dr. I. Melzer

Dept. of Physical Therapy, Recanati School for Community Health Professions, Faculty of Health Sciences, Ben-Gurion University of the Negev, Beer Sheva 84105, Israel

Phone: (972-8) 647-7735

Fax: (972-8) 647-7683

email: itzikm@bgu.ac.il

References

1. Bergen G, Stevens MR, Burns ER. Falls and fall injuries among adults aged ≥ 65 years—United States, 2014. *MMWR Morb Mort Wkly Rep* 2016; 65: 993-8.
2. Aizen E, Nixon H, Shugaev I. Awareness and functional outcome of hip fracture-related falls among patients with a history of recurrent falling. *IMAJ* 2018; 20 (1): 38-42
3. Cumming RG, Klineberg RJ. Fall frequency and characteristics and the risk of hip fractures. *J Am Geriatr Soc* 1994; 42 (7): 774-8.
4. Glaister BC, Bernatz GC, Klute GK, Orendurff MS. Video task analysis of turning during activities of daily living. *Gait Posture* 2007; 25 (2): 289-94.
5. Dite W, Temple VA. Development of a clinical measure of turning for older adults. *Am J Phys Med Rehabil* 2002; 81 (11): 857-66; quiz 867-8.
6. Thigpen MT, Light KE, Creel GL, Flynn SM. Turning difficulty characteristics of adults aged 65 years or older. *Phys Ther* 2000; 80 (12): 1174-87.
7. Booth ML. Assessment of physical activity: an international perspective. *Res Q Exerc Sport*; 71 (2 Suppl): S114-20.
8. Fleiss JL. *The Design and Analysis of Clinical Experiments*. New York: Wiley, 1986.

9. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1 (8476): 307-10.
10. Monti KL. Folded empirical distribution function curves – mountain plots. *Am Stat* 1995; 49: 342-5.
11. Berg KO, Maki BE, Williams JI, Holliday PJ, Wood-Dauphinee SL. Clinical and laboratory measures of postural balance in an elderly population. *Arch Phys Med Rehabil* 1992; 73 (11): 1073-80.
12. Rockwood K, Awalt E, Carver D, MacKnight C. Feasibility and measurement properties of the functional reach and the timed up and go tests in the Canadian study of health and aging. *J Gerontol A Biol Sci Med Sci* 2000; 55 (2): M70-3.
13. Steffen TM, Hacker TA, Mollinger L. Age- and gender-related test performance in community-dwelling elderly people: Six-Minute Walk Test, Berg Balance Scale, Timed Up & Go Test, and gait speeds. *Phys Ther* 2002; 82 (2): 128-37.
14. Faber MJ, Bosscher RJ, van Wieringen PC. Clinimetric properties of the performance-oriented mobility assessment. *Phys Ther* 2006; 86 (7): 944-54.
15. Nevitt MC, Cummings SR, Kidd S, Black D. Risk factors for recurrent nonsyncopal falls. A prospective study. *JAMA* 1989; 261 (18): 2663-8.
16. Bogle Thorbahn LD, Newton RA. Use of the Berg Balance Test to predict falls in elderly persons. *Phys Ther* 1996; 76 (6): 576-83; discussion 584-5.
17. Shumway-Cook A, Baldwin M, Polissar NL, Gruber W. Predicting the probability for falls in community-dwelling older adults. *Phys Ther* 1997; 77 (8): 812-19.
18. Bhatt T, Espy D, Yang F, Pai YC. Dynamic gait stability, clinical correlates, and prognosis of falls among community-dwelling older adults. *Arch Phys Med Rehabil* 2011; 92 (5): 799-805.
19. Lin MR, Hwang HF, Hu MH, Wu HD, Wang YW, Huang FC. Psychometric comparisons of the timed up and go, one-leg stand, functional reach, and Tinetti balance measures in community-dwelling older people. *J Am Geriatr Soc* 2004; 52 (8): 1343-8.
20. Harada N, Chiu V, Damron-Rodriguez J, Fowler E, Siu A, Reuben DB. Screening for balance and mobility impairment in elderly individuals living in residential care facilities. *Phys Ther* 1995; 75 (6): 462-9.

Capsule

Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation

Several studies have investigated links between the gut microbiome and colorectal cancer (CRC), but questions remain about the replicability of biomarkers across cohorts and populations. **Maltez Thomas** and co-authors performed a meta-analysis of five publicly available datasets and two new cohorts and validated the findings on two additional cohorts, considering in total 969 fecal metagenomes. Unlike microbiome shifts associated with gastrointestinal syndromes, the gut microbiome in CRC showed reproducibly higher richness than controls ($P < 0.01$). This finding is partially due to expansions of species typically derived from the oral cavity. Meta-analysis of the microbiome functional potential identified gluconeogenesis and the putrefaction and fermentation pathways as being associated with CRC, whereas the stachyose and starch degra-

dation pathways were associated with controls. Predictive microbiome signatures for CRC trained on multiple datasets showed consistently high accuracy in datasets not considered for model training and independent validation cohorts (average area under the curve, 0.84). Pooled analysis of raw metagenomes showed that the choline trimethylamine-lyase gene was overabundant in CRC ($P = 0.001$), identifying a relationship between microbiome choline metabolism and CRC. The combined analysis of heterogeneous CRC cohorts thus identified reproducible microbiome biomarkers and accurate disease-predictive models that can form the basis for clinical prognostic tests and hypothesis-driven mechanistic studies.

Nature Med 2019; 25: 667

Eitan Israeli

Capsule

Structural variation in the gut microbiome associates with host health

Differences in the presence of even a few genes between otherwise identical bacterial strains may result in critical phenotypic differences. **Zeevi** and co-workers systematically identified microbial genomic structural variants (SVs) and found them to be prevalent in the human gut microbiome across phyla and able to replicate in different cohorts. SVs are enriched for CRISPR-associated and antibiotic-producing functions and depleted from housekeeping genes, suggesting that they have a role in microbial adaptation. The authors found multiple associations between SVs and host disease risk factors, many of which replicate in an independent cohort.

Exploring genes that are clustered in the same SV, they uncovered several possible mechanistic links between the microbiome and its host, including a region in *Anaerostipes hadrus* that encodes a composite inositol catabolism-butyrate biosynthesis pathway, the presence of which is associated with lower host metabolic disease risk. Overall, these results uncover a nascent layer of variability in the microbiome that is associated with microbial adaptation and host health.

Nature 2019; 568: 43

Eitan Israeli

“A teacher is one who makes himself progressively unnecessary”

Thomas Neely Carruthers (1900–1960), U.S. bishop of the Episcopal Diocese of South Carolina