American Journal of Emergency Medicine 37 (2019) 457–460

Contents lists available at ScienceDirect

American Journal of Emergency Medicine

journal homepage: www.elsevier.com/locate/ajem

Original Contribution

Effectiveness of the "Timed Up and Go" (TUG) and the Chair test as screening tools for geriatric fall risk assessment in the ED☆

Richard B. Chow, DO [a], Andre Lee, BS, MS [a], Bryan G. Kane, MD [a], Jeanne L. Jacoby, MD [a], Robert D. Barraco, MD, MPH [b], Stephen W. Dusza, DrPH [a], Matthew C. Meyers, BS [a], Marna Rayl Greenberg, DO, MPH [a,]*

[a] Department of Emergency and Hospital Medicine, Lehigh Valley Hospital and Health Network/USF MCOM, Allentown, PA, United States
[b] Department of Surgery, Lehigh Valley Hospital and Health Network/USF MCOM, Allentown, PA, United States

ARTICLE INFO

Article history:
Received 5 May 2018
Received in revised form 2 June 2018
Accepted 6 June 2018

Keywords:
Sex differences
Elderly falls
TUG test
Chair test

ABSTRACT

Objective: We sought to evaluate the effectiveness of the "Timed Up and Go" (TUG) and the Chair test as screening tools in the Emergency Department (ED), stratified by sex.
Methods: This prospective cohort study was conducted at a Level 1 Trauma center. After consent, subjects performed the TUG and the Chair test. Subjects were contacted for phone follow-up and asked to self-report interim falling.
Results: Data from 192 subjects were analyzed. At baseline, 71.4% (n = 137) screened positive for increased falls risk based on the TUG evaluation, and 77.1% (n = 148) scored below average on the Chair test. There were no differences by patient sex.
By the six-month evaluation 51 (26.6%) study participants reported at least one fall. Females reported a non-significant higher prevalence of falls compared to males (29.7% versus 22.2%, p = 0.24). TUG test had a sensitivity of 70.6% (95% CI: 56.2%–82.5%), a specificity of 28.4% (95% CI: 21.1%–36.6%), a positive predictive (PP) value 26.3% (95% CI: 19.1%–34.5%) and a negative predictive (NP) value of 72.7% (95% CI: 59.0%–83.9%). Similar results were observed with the Chair test. It had a sensitivity of 78.4% (95% CI: 64.7%–88.7%), a specificity of 23.4% (95% CI: 16.7%–31.3%), a PP value 27.0% (95% CI: 20.1%–34.9%) and a NP value of 75.0% (95% CI: 59.7%–86.8%). No significant differences were observed between sexes.
Conclusions: There were no sex specific significant differences in TUG or Chair test screening performance. Neither test performed well as a screening tool for future falls in the elderly in the ED setting.

1. Introduction

Falls remain one of the top 20 most expensive medical conditions with hospital costs averaging over $30,000 per encounter; in 2015, alone, the costs for falls for those aged 65 and above to Medicare totaled $31 billion [1]. Recent research efforts of fall interventions in older adults have identified many differences in risk factors associated with sex in determining geriatric fall risks—in one example, a recent longitudinal study in Age and Aging found that age, depressive symptoms, and performance on standing balance tests were separate determinants for men, while incontinence and frailty increased fall risks in women [2]. Another cross-sectional study found unique associations to each sex such as nutritional risks, osteoporosis and even levels of household income [3].

Contrary to these reported differences, fall risks are clinically assessed the in the same manner for both sexes. These assessments tend to be functional in nature despite the limited success in predicting fall risk [4]. A recent systematic review identified a protective bias towards women needing fall prevention programs more than men (though increased activity is protective for both males and females), while males consistently sustain higher death rates and Disability-Adjusted Life Years (DALYs) lost worldwide [5]. More research is

☆ The authors have no outside support information, conflicts or financial interest to disclose. This study, in part, was funded by an unrestricted community grant from the Ann and Carl Anderson Trust and this work has been presented, in part, as an abstract at the 2017 IGM Congress in Sendai, Japan.Authors would like to acknowledge the research operations management of Anita Kurt, PhD, RN, the team member assistance for screening, enrolling, coordinating, and or phone follow-up of Tyler M. Adams, BS, Steven J. Berk, DO, Vartika Bhardwaj, MD, Tracy M. Bishop, DO, Ryan Day, MD, Danielle Mills, MD, Bernadette Glenn-Porter, BS, Shaye M. Glovas, BS, Victoria Goodheart, DO, Michael Goodwin, MD, Kyli N. Krape, DO, Stephanie S. Merrick, MD, Kayley J. Miller, MSPAS, PA-C, Elizabeth C. Moore, DO, Sofia M. Murillo, BS, Samantha L. Myles BSN, RN, Cameron Paterson, MD, Todd Remaley, DO, Rolando E. Rios, MD, Lauren E. Semler, MS, MSPAS, PA-C, Ryan M. Surmaitis, DO, Kara Mia Villanueva, MS, Michael Wagner, DO, Deirdre Warner, DO, Adison Weseloh, MD, Alexander Youngdahl, DO, and Phillip Zegelbone, MD.
* Corresponding author at: 1909 Earls Court, Allentown, PA 18103, United States.
E-mail address: mrgdo@ptd.net (M.R. Greenberg).

https://doi.org/10.1016/j.ajem.2018.06.015
0735-6757/

required to identify and validate sex differences in fall risk, so that practitioners may be better equipped to assess them.

Serving as the frontline of acute healthcare, the Emergency Department (ED) may be an ideal place to recognize these factors and initiate a pathway for early multidisciplinary interventions. Already, unintentional falls account for about 13.5% of Emergency Department (ED) visits, a staggering 3 million emergency department visits annually [6, 7]. There is a dearth of studies that seek to identify those at risk of falls during an encounter, including for non-traumatic visits in the ED. There has been a call for streamlined screening protocols in the ED for the geriatric population [8].

The Timed Up and Go Test (TUG) and the 30-Second Chair test are proposed by the CDC [9, 10] as a means of screening elderly patients for increased risk for falls. While the TUG test has shown a decreased performance in females compared to males among the geriatric mild cognitive impairment (MCI) population [11], the Chair test has not been evaluated for effectiveness by sex. Neither test has been evaluated in the ED setting. In this setting, a desirable test would offer an objective means of screening while being effortless enough to be performed within the time constrained environment. Such a tool would optimally assist clinicians in mitigating future morbidity/mortality of their elderly patients. In this study, we sought to evaluate the effectiveness of the TUG and Chair test in the ED setting, stratified by sex.

2. Methods

After hospital IRB review and approval, this prospective trial was conducted in the ED of a Level 1 Trauma Center in Northeastern Pennsylvania with an annual census of over 90,000 adult visits per year. Participants were screened and included in the study if they were ≥65 years old, being discharged from the ED, English speaking, had capacity for consent, and personally identified a risk factor for falling. Risk factors for falling were based on Centers for Disease Control (CDC) guidelines [12] and were deemed positive (study inclusive) if a patient reported that they had either fallen in the last year, worried about falling, or admitted that they felt unsteady when standing or walking. After obtaining informed consent, patients were evaluated with the TUG and Chair test.

The TUG test was administered with the participant initially seated in a chair. The participant was then instructed to stand up from the seated position, walk in a straight line to a marked point 10 ft from the chair, turn and walk back to the chair and return to a seated position. Study personnel timed the participant with a stopwatch, and recorded in seconds, the time it took from the point of directing the participant to begin the movement until they returned to the seated position. Participants requiring ≥12 s to complete the TUG test were identified as at risk for fall as defined by the CDC [9].

The 30-s Chair test was also administered with the participant initially seated in a chair. The chair utilized for study testing was a standard 17″ high chair without armrests. Subjects were instructed to sit in the middle of the chair, and to place hands on the opposite shoulder crossed at the wrists. With feet flat on the floor, subjects were instructed to keep their back straight and arms against chest, and rise to a full standing position and to return to a seated position, and to repeat this exercise as many times as possible in 30 s. The total number of completed cycles of this exercise was counted and recorded by study personnel. Subject performance was then evaluated by their age and gender group standards as defined by the CDC [10] as follows:

Age group	Men	Women
60–64	<14	<12
65–69	<12	<11
70–74	<12	<10
75–79	<11	<10
80–84	<10	<19
85–89	<8	<18
90–94	<7	<4

For analysis participants were considered “below average”, or “at or above average” if they did not meet, or if they exceeded these criteria, respectively.

In addition to the TUG and Chair tests, questionnaires were administered to the participants to quantify falls risk and to collect information of covariates that can be associated with falls. Questions were asked related to overall health, the use of assistive devices to ambulate, and to quantify the number of falls experienced in the year prior to study participation.

All study participants were actively followed up by phone for 6 months after their initial ED visit and enrollment in the study. During this time period, participants were contacted, and were asked to self-report any falls that had occurred in the months since enrollment. Patients in the data set who reported falling more than once during the prescribed 6-month follow up period were counted once.

Descriptive statistics and graphical methods were used to describe the study population. Relative frequencies and chi-square tests were used to assess the distributions of participant characteristics by participant sex. The primary outcome measure was a self-reported fall that occurred between the baseline and six-month evaluation. The outcome was a dichotomous measure. The two primary independent variables were the assessments from the TUG test and Chair tests collected at baseline. Measures of diagnostic performance of these screening exams were estimated along with 95% confidence intervals. General linear models with a log-link were used to assess any differences in sensitivity and specificity measures by participant sex. These models used reported falls by six months as the outcome and gender as the main covariate. The association between a self-reported fall at 6-months and the results of the screening tests were performed using logistic regression. Additional covariates such as age, sex, falls in the year prior to the study, self-assessed health status and use of an assistive device were included in the models to evaluate potential confounding. Additionally the performance of both Tug and Chair tests was entered into models continuously in terms of an ROC analysis for risk of future falls —this was to insure the predictive value of the tests in case the test was impacted by published thresholds that weren't calibrated for an ED population. (Appendix) All analyses were performed using Stata v14.2. (Stata Corporation, College Station, TX).

3. Results

Two hundred participants were enrolled into this study. Of these, 8 were excluded from analysis because they: withdrew consent ($n = 3$), were lost to follow-up ($n = 2$), died prior to 6-month follow-up evaluation ($n = 2$), or were deemed ineligible after enrollment ($n = 1$) (See Fig. 1). The resultant sample size was 192 participants. The characteristics of the study sample, stratified by participant gender, are presented in Table 1. A majority of the participants were female ($n = 111$, 57.8%). The average age for the sample was 74.4 years ($SD = 7.4$), with no differences in age between genders, $p = 0.34$. Only 37.0% of the sample assessed their own health as very good or excellent. Over 69% of the study participants (69.1% male and 69.4% female, $p = 0.97$) reported a fall in the previous year. At the baseline evaluation, 71.4% ($n = 137$) of the sample screened positive for increased falls risk based on the TUG evaluation, and 77.1% ($n = 148$) scored below average on the Chair test. Results of the evaluations were not different by patient sex.

By the six-month evaluation 51 (26.6%) study participants reported at least one fall since the baseline evaluation. Females reported a non-significant higher prevalence of falls compared to males (29.7% versus 22.2%, $p = 0.24$). Of these 51 participants who reported a fall, 36 (70.6%) failed the TUG test during the baseline evaluation. However, 101 of the 141 participants (71.6%) who did not report a fall failed the TUG test at baseline, resulting in a sensitivity of 70.6% (95% CI: 56.2%–82.5%), a specificity of 28.4% (95% CI: 21.1%–36.6%), a positive predictive value 26.3% (95% CI:19.1%–34.5%) and a negative predictive value of 72.7% (95% CI: 59.0%–83.9%). Similar results were observed with the

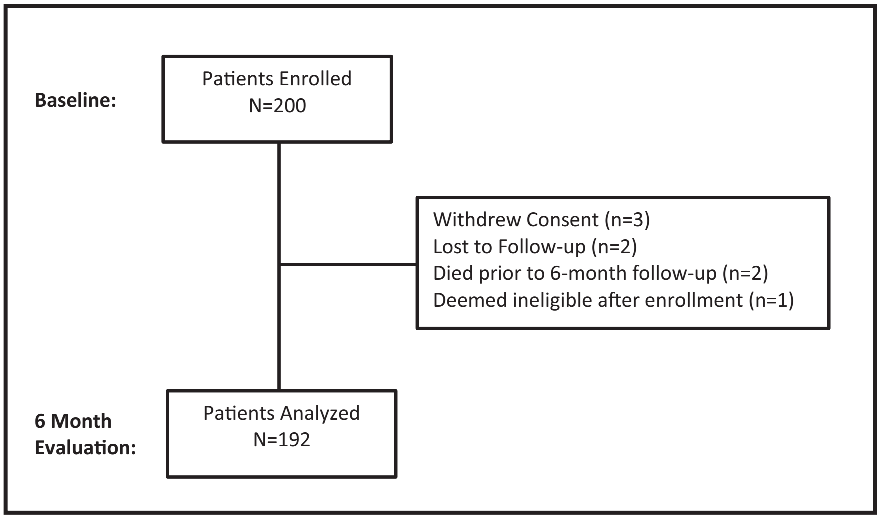

Fig. 1. Schematic of study population.

Chair test. Of the 51 participants who reported a fall, 40 (78.4%) failed the Chair test at the baseline evaluation, resulting in in a sensitivity of 78.4% (95% CI: 64.7%–88.7%), a specificity of 23.4% (95% CI: 16.7%–31.3%), a positive predictive value 27.0% (95% CI: 20.1%–34.9%) and a negative predictive value of 75.0% (95% CI: 59.7%–86.8%).

The ROC curve for the performance of the Tug and Chair Test were both poor, ($Area_{ROC} = 0.54$ and $Area_{ROC} = 0.53$ respectively), detailed results can be found in the Appendix. Sensitivity and specificity and the percent correctly classified (to assess for potential cutoff points) can also be found in the Appendix.

No significant differences were observed in screening test performance between male and female participants. The differences in sensitivity and specificity for the TUG test observed between male and female participants were + 11.1% (p = 0.38) and −2.5% (p = 0.74), respectively. The differences in sensitivity and specificity for the Chair test observed between male and female participants were −1.1% (p = 0.93) and −10.7% (0.13), respectively. The only factor that was observed to be associated with a reported fall by the six-month evaluation was a positive reported personal history of falling in the year prior to the baseline evaluation. Participants who reported a fall in the year prior to the baseline evaluation were 2.6 times more likely to report a fall at the six-month follow-up than those who did not (OR = 2.6; 95% CI:1.2–5.7).

Table 1
Baseline characteristics of the study population, stratified by participant gender (n = 192).

	Overall	Male	Female	p-Value
	n = 192	n = 81	n = 111	
Age, mean (SD)	74.4 (7.4)	75.0 (7.9)	74.0 (7.1)	0.34
Self-assessed health				
Poor	6 (3.1)	2 (2.5)	4 (3.6)	0.88
Fair	33 (17.2)	16 (19.8)	17 (15.3)	
Good	82 (42.7)	35 (43.2)	47 (42.3)	
Very good	51 (26.6)	21 (25.9)	30 (27.0)	
Excellent	20 (10.4)	7 (8.6)	13 (11.7)	
Use assistive device regularly				
No	138 (71.9)	60 (74.1)	78 (70.3)	0.56
Yes	54 (28.1)	21 (25.9)	33 (29.7)	
Fall in past year				
No	59 (30.7)	25 (30.9)	34 (30.6)	0.97
Yes	133 (69.3)	56 (69.1)	77 (69.4)	
Tug test				
Elevated falls risk	137 (71.4)	60 (74.1)	77 (69.4)	0.48
Normal falls risk	55 (28.7)	21 (25.9)	34 (30.6)	
Chair test				
Below average	148 (77.1)	66 (81.5)	82 (73.9)	0.22
At or above average	44 (22.9)	15 (18.5)	29 (26.1)	

When looking at these results stratified by participant gender, the effect is slightly attenuated in females (OR = 2.0; 95% CI: 0.8–5.1), and modestly larger in males (OR = 4.6; 95% CI: 1.0–21.8). The diagnostic accuracy and predictive measures for TUG and Chair test, stratified by participant gender can be found in Table 2.

4. Discussion

In 2014, at a Society for Academic Emergency Medicine consensus conference, it was determined that a prioritized research agenda should include finding the most feasible falls-risk tool for men and women in the ED setting [13]. Our study indicates that the reliance on using TUG and Chair tests as a means of predicting fall risk shows limited predictive ability. In our study, simply asking a patient if they have fallen in the past year was found to be a better screening tool. CDC recommendations on the evaluation of falls are predicated on considerations other than the TUG and Chair tests. The other considerations associated with each gender falling include environmental factors related to age [2, 3]. A more comprehensive approach might include the use of other scoring methods and evaluative tools in the determination of future falls, as demonstrated in a retrospective cohort study that used a multifactorial questionnaire for a more comprehensive assessment [4, 14].

There have been studies that have attempted to delve into the reliability of the original TUG reports. A recent prospective study by Kojima correlated increasing TUG times with falls among a small group of 259 British elderly patients [15]. That study denoted a high specificity and specifically correlated a negative predictive value of falls for individuals who did well with the TUG test. By comparison, our results indicate a poor association. Of note is that the study populations were recruited from community dwellers as opposed to an ED setting.

The utility of using TUG in reference to other aspects of elderly function is much more varied however. A Canadian study by Rockwood specified what they felt was poor test - retest reliability of the TUG test [16]. However, the study measured outcomes in comparison of those who were cognitively impaired versus those who were not in relation to ability to complete activities of daily living instead of fall risk [16]. A systematic review has also brought into question the utility of the TUG test [17]. Other studies seem to support the utility of using the TUG test in quantifying functional mobility as a function of TUG scoring as noted by Podsiadlo [18].

Another issue with the TUG test has been the lack of an established international cut-off that differed from the CDC's standard of a positive result at 12 s [14]. Its use in various studies around the world reports a cutoff of 12.6 s in England, 13.5 s in Japan, and 14.5 in Spain [15, 19, 20].

Table 2
Diagnostic accuracy and predictive measures for TUG and CHAIR test, stratified by participant sex.

Group	Exam	Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio
Overall	TUG Test	70.6%	28.4%	26.3%	72.7%	0.98	1.04
	Chair Test	78.4%	23.4%	27.0%	75.0%	1.02	0.92
Female	TUG Test	66.7%	29.5%	28.6%	67.6%	0.95	1.10
	Chair Test	78.8%	28.2%	31.7%	75.9%	1.10	0.75
Male	TUG Test	77.8%	27.0%	23.3%	81.0%	1.07	0.82
	Chair Test	77.8%	17.5%	21.2%	73.3%	0.94	1.27

It was also used as a diagnostic tool in Brazil for normal pressure hydrocephalus (NPH) using a cutoff of 16.5 s, with excellent specificity (0.967) and sensitivity (0.933) [21]. This variability in geographical standards and uses suggests a possibility that our cutoff was not discriminatory for falls because we have the wrong 'set point'. However in our detailed analysis, there did not seem to be a better set point to report. Further analysis at different set points may be warranted in future studies with a more robust sample. From a public health perspective, having a predictive and efficient screening tool for use in the ED setting could have a tremendous effect. While the tools studied herein are ineffective in the population studied, further research aimed at finding a method that is successful is vital.

4.1. Limitations

There are several limitations to be considered in the generalized applicability of this study. First, the study population screened included only elderly individuals who were already at risk of falling (as defined by the CDC) [12]. A possible opportunity exists to expand this study to be inclusive of all elderly individuals aged 65 and older without such restrictions. It is conceivable that the tests would have improved performance when applied to an overall ED older adult population at lower risk of falling. Further, this was a convenience sample of patients who were enrolled when a study member was present to screen and obtain consent. Given that there was likely a lack of such members during night shifts this may also impact the study by excluding patients who are likely to fall (given adequate lighting is likely a contributor for falls). The study also excluded those individuals who did not speak English. Other exclusion criteria were for those who required admission to the hospital. Further studies could serve to enroll and follow up with such patients after discharge from the hospital.

A final limitation needs considered. The baseline reported fall rate of the population as studied was 69% and yet the self-reported fall rate at 6 months was 26%, this suggests that either falls before or after the ED visits are being reported differently, or that the ED visits and the study protocol exerted a significant protective effect on future falls. Future study could be aimed at improving the potential for reporting bias or corroborating our findings.

5. Conclusions

In our study, the TUG and Chair tests, applied at the given thresholds, do not add significant additional screening performance when added to the screen already applied by the inclusion criteria (patients reporting that they had either fallen in the last year, worried about falling, or admitted that they felt unsteady when standing or walking). Additionally, there were no sex specific significant differences in TUG or Chair test screening performance. Further research to determine what might be the best tool to predict fall risk in this setting is needed.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ajem.2018.06.015.

References

[1] Burns EB, Stevens JA, Lee RL. The direct costs of fatal and non-fatal falls among older adults—United States. J Safety Res 2016;58:99–103.
[2] Gale CR, Cooper C, Aihie Sayer A. Prevalence and risk factors for falls in older men and women: the English longitudinal study of ageing. Age Ageing 2016;45(6):789–94.
[3] Chang VC, Do MT. Risk factors for falls among seniors: implications of gender. Am J Epidemiol 2015 Apr 1;181(7):521–31.
[4] Ibrahim A, Singh DKA, Shahar S, Omar MA. Timed up and go test combined with self-rated multifactorial questionnaire on falls risk and sociodemographic factors predicts falls among community-dwelling older adults better than the timed up and go test on its own. J Multidiscip Healthc 2017 Oct 26;10:409–16.
[5] Sandlund M, Skelton DA, Pohl P, Ahlgren C, Melander-Wikman A, Lundin-Olsson L. Gender perspectives on views and preferences of older people on exercise to prevent falls: a systematic mixed studies review. BMC Geriatr 2017 Feb 17;17(1):58.
[6] Centers for Disease Control. Take a stand on falls. https://www.cdc.gov/features/older-adult-falls/index.html. [Accessed 02/23/2018].
[7] Gillespie LD, Gillespie WJ, Robertson MC, Lamb SE, Cumming RG, Rowe BH. Interventions for preventing falls in elderly people. Cochrane Database Syst Rev 2003(4) 2009;(2):CD000340. (Update in: Cochrane Database Syst Rev.).
[8] Carpenter CR, Avidan MS, Wildes T, Stark S, Fowler SA, Lo AX. Predicting geriatric falls following an episode of emergency department care: a systematic review. Acad Emerg Med 2014 Oct;21(10):1069–82.
[9] Centers for Disease Control. Assessment, "Timed up and go.". https://www.cdc.gov/steadi/pdf/TUG_Test-print.pdf. [Accessed 02/23/2018].
[10] Centers for Disease Control. Assessment, "30-second chair stand.". https://www.cdc.gov/steadi/pdf/30_Second_Chair_Stand_Test-print.pdf. [Accessed 02/23/2018].
[11] Ibrahim A, Singh DKA, Shahar S. "Timed up and go" test: age, gender and cognitive impairment stratified normative values of older adults. Ginsberg SD. PLoS ONE 2017;12(10):e0185641.
[12] Centers for Disease Control. Algorithm for fall risk screening, assessment, and intervention. https://www.cdc.gov/steadi/pdf/STEADI-Algorithm-508.pdf. [Accessed 02/23/2018].
[13] Greenberg MR, Kane BG, Totten VY, Raukar NP, Moore EC, Sanson T, et al. Injury due to mechanical falls: future directions in gender-specific surveillance, screening, and interventions in emergency department patients. Acad Emerg Med 2014 Dec;21(12):1380–5.
[14] Beauchet O, Fantino B, Allali G, Muir SW, Montero-Odasso M, Annweiler C. Timed up and go test and risk of falls in older adults: a systematic review. J Nutr Health Aging 2011 Dec;15(10):933–8.
[15] Kojima G, et al. Does the timed up and go test predict future falls among British community-dwelling older people? Prospective cohort study nested within a randomised controlled trial. BMC Geriatr 2015;15:38.
[16] Rockwood K, Awalt E, Carver D, Macknight C. Feasibility and measurement properties of the functional reacn and the timed up and go tests in the Canadian study of health and aging. J Gerontol A Biol Med Sci 2000;55A:M70–3.
[17] 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.
[18] Podsiadlo D, Richardson S. The timed up & go: a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991;39:142–8.
[19] Zakaria NA, Kuwae Y, Tamura T, Minato K, Kanaya S. Quantitative analysis of fall risk using TUG test. Comput Methods Biomech Biomed Engin 2015;18(4):426–37.
[20] Del-Río-Valeiras M, Gayoso-Diz P, Santos-Pérez S, Rossi-Izquierdo M, Faraldo-García A, Vaamonde-Sánchez-Andrade I, ... Soto-Varela A. Is there a relationship between short FES-I test scores and objective assessment of balance in the older people with age-induced instability? Arch Gerontol Geriatr 2016:6290–6.
[21] Mendes GS, de Oliveira MF, Pinto FG. The timed up and go test as a diagnostic criterion in normal pressure hydrocephalus. World Neurosurg 2017;105456-461.