Reliability Test of The Timed Up and Go Test in Elderly People with Knee Osteoarthritis

Dwi Rosella Komala Sari, Annisa Firsita Motik

Department of Physiotherapy, Faculty of Health Sciences, Universitas Muhammadiyah Surakarta

Correspondence email: drks133@ums.ac.id

Submition: 2024-01-26; Accepted: 2024-03-06; Published: 2024-06-01

ABSTRACT

Osteoarthritis (OA) of the knee is a degenerative disease in elderly. Pain is a common problem that can decrease balance ability in static and dynamic balance. The Timed Up and Go test (TUG) is one of the tests that could be used to evaluate dynamic balance. However, the reliability of this test in knee OA has not been proven in Indonesia. Therefore, this study aimed to determine the reliability of TUG in elderly with OA knee, especially for test-retest and inter-rater reliability. This study was an observational study method with a methodological research approach. There were 58 respondents involved in this study who were taken by purposive sampling and met the inclusion and exclusion criteria. Analysis test revealed test-retest reliability and inter-rater reliability have ICC 0.982 and 0.995 (excellent). The test-retest has $\alpha=0.991$ (excellent), $p<0.001$, and $r=0.962$ (very strong correlation). Meanwhile, inter-rater reliability was obtained $\alpha=0.997$ (excellent), $p<0.001$, and $r=0.990$ (very strong correlation). Therefore, the TUG test is reliable for test-retest and inter rater as a measure of dynamic balance as well recommended for elderly with knee osteoarthritis.

Keywords: reliability, timed up and go test, knee osteoarthritis

INTRODUCTION

The physical abilities declination in the elderly is caused by degenerative processes, one of which is knee osteoarthritis (OA) (Martel-Pelletier et al., 2016). Knee OA is known as a degenerative disease of the knee joint, usually caused by progressive wear and tear, and loss of articular cartilage. This condition most commonly occurs in the elderly. The knee consists of bone structures (distal femur, proximal tibia and patella), cartilage (hyaline cartilage and meniscus), ligaments and synovial membrane. The synovial membrane is responsible for providing lubrication and nutrition to the articular cartilage with synovial fluid. High use and pressure on the joints can cause pain including Osteoarthritis (OA) (Mora Herrera et al., 2018).

OA is classified into two groups, namely primary (idiopathic or non-traumatic) and secondary (usually caused by trauma or mechanical misalignment). The severity of the disease can be assessed using radiology and using the Kellgren–Lawrence (KL) classification in 1957. OA was believed to be a degenerative disease of the cartilage, but recent evidence proves that there are several causal factors involved such as mechanical forces, inflammation, trauma, biochemical reactions, and metabolic diseases (Mora Herrera et al., 2018) such as obesity, diabetes and insulin resistance, dyslipidemia, and hypertension (Courties et al., 2017).

According to (Riskesdas, 2018), OA is a joint disease that has a prevalence of 7.3%, with women at 8.46% and men at 6.13%. Prevalence increases in proportion to age level. Age group 15-24 years was 1.23%, 25-35 years 3.10%, 35-44 years 6.27%, 45-54 years 11.08%, 55-64 years 15.55%, 65- 74 years 18.63%, and >75 years 18.95%.

Pain is a common problem in knee OA that can reduce some physical performance, like dynamic balance ability. Postural stability is
reduced due to joint laxity, obesity, as well as neuromuscular changes such as muscle weakness and low proprioception that occur during the aging process (Hassan et al., 2001). The Timed Up and Go test (TUG) is a measuring tool used to evaluate and identify mobility and dynamic balance disorders in the elderly. Assessment is one of the things that must be done as an initial screening so that if there is a balance disorder, early efforts can be made to prevent unwanted injuries. So, before being applied, measuring instruments must be tested as valid and reliable. If it has been proven to be reliable, accurate and consistent result data will be obtained, so it can be used as an evaluation parameter. However, the reliability of TUG for knee OA conditions in Indonesia has never been reported in previous research, therefore TUG still needs to be proven and its reliability observed more deeply. So this research aims to determine the reliability of the TUG measuring instrument in terms of test-retest and inter-rater in elderly people with knee OA.

MATERIALS AND METHODS
This research was carried out based on a research permit from the Research Ethics Committee of the Faculty of Medicine, Muhammadiyah University of Surakarta with number 4538/B.1/KEPK-FKUMS/X/2022. The research design used an observational study type with a methodological research approach, to determine the reliability of the TUG as a tool for measuring dynamic balance in individuals with knee OA. This research was conducted at the Rumah Sakit Islam Klaten from December 2022 to January 2023. We analyzed the internal consistency, reliability consistency, the correlation between test 1 and test 2 and between the measurement results of rater 1 and rater 2, as well as the correlation coefficient on test-retest reliability and inter-rater reliability. Meanwhile, the independent variable is dynamic balance.

The sampling technique used purposive sampling carried out by screening first through a sample screening process with the inclusion criteria and exclusion criteria. The inclusion criteria included (1) having signs and symptoms of OA according to the American College of Rheumatology (ACR), (2) unilateral or bilateral knee OA, (2) age ≥ 45 years, (3) diagnosis of knee OA by X-ray and identified using the Kellgren & Lawrence classification for grade ≥2, (4) having knee pain >3 using the Numeric Rating Scale, (5) able walk without using assistive devices. Meanwhile, the exclusion criteria involved (1) received intraarticular corticosteroid injection, (2) the patient took steroid medication, (3) had a history of surgery on the lower extremities, (4) had neurological diseases such as stroke, Parkinson's, cerebral ataxia, (5) had complaints of low back pain which is related to knee pain, (6) and the patient is unable to communicate well.

The sample size calculation used sample size calculation software based on the formula from (Arifin, 2022) with a minimum acceptable of 0.6, expected reliability of 0.8, significance level of 0.05, power of 80%, and dropout of 10%. Then, the sample size in this study was 58 people.

There are 2 types of reliability tests in this study, e.g.; test-retest and inter-rater respectively. Inter-rater measurements were carried out first by rater 1 and rater 2 at the same time. When the ‘go’ instruction is said, the respondent stands up from the chair and starts walking 3 meters to the line marked by the cone, then circles the cone and walks towards the chair again until he sits down. Rater 1 and Rater 2 counted the time using a stopwatch from when ‘go’ was said until they sat back down in the chair. After 15 minutes, a retest was carried out by rater 1 with the same test procedures and time measurements using a stopwatch from starting standing, walking 3 meters around the cone until sitting back on the chair.

The data analysis used in this research is (1) univariate test, as well presenting data in frequency form for one variable (Notoatmodjo, 2010) which describes the characteristics of respondents, in the form of minimum value, maximum value, frequency (n), average (mean) and standard deviation (SD), (2) normality test on the measurement results of rater 1 test 1, rater 1 test 2 and on rater 2, (3) reliability test by displaying internal consistency results (Cronbach's Alpha) with an interpretation of ≥0.9 (very good/excellent), ≥0.8 (good), ≥0.7 (acceptable), ≥0.6 (questionable), ≥0.5 (not good or weak), and <0.5 (unacceptable) (Mallery, 2019). Reliability consistency values (Intra Class Correlation/ICC) interpreted by <0.5 (low
reliability), 0.5–0.75 (medium reliability), 0.75–
0.90 high reliability, and >0.90 very high
reliability (excellent) (Koo & Li, 2016). (4)
correlation test between the results of test 1 and
test 2 on test-retest reliability and between the
results of rater 1 and rater 2 on inter-rater
reliability with the interpretation of a p-value
≤0.001 meaning very high significance, p <0.01
very significant, p <0.05 significant, p <0.1 is not
significant, and p ≥0.1 is not very significant
(Raiola & Di tore, 2012). For the correlation
coefficient used interpretation value as 0.01 –
0.09 as meaningless relationship, 0.10 – 0.29 for
a weak relationship, 0.30 – 0.49 for a moderate
relationship, 0.50 – 0.69 for a strong relationship,
0.70 – 0.85 for a very strong relationship, >0.90
perfect relationship (Cicchetti, 1994).

RESULTS

Table 1 explains that the average age is
above 60 years, with a minimum age of 48 years
and a maximum of 81 years. The female gender is
dominant in this study and the bilateral side was
dominant. The average duration of suffering from
knee OA of 42 months or 3.5 years. Dynamic
balance (TUG) was 16.8 seconds, which is
categorized as low dynamic balance or has a high
risk of falling. Because more than the value of
13.5 seconds is a reference value for high
dynamic balance (Shumway-Cook et al., 2000).

<table>
<thead>
<tr>
<th>Characteristics Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Table 1. Characteristics of research data</strong></td>
</tr>
<tr>
<td><strong>Variable</strong></td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>Men</td>
</tr>
<tr>
<td>Women</td>
</tr>
<tr>
<td>OA Status</td>
</tr>
<tr>
<td>Unilateral</td>
</tr>
<tr>
<td>Bilateral</td>
</tr>
<tr>
<td>OA Duration (month)</td>
</tr>
<tr>
<td>TUG (sec)</td>
</tr>
<tr>
<td>OA Grade (Kallgren and Lawrence)</td>
</tr>
<tr>
<td>Right</td>
</tr>
<tr>
<td>Left</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>3.2 Test-retest reliability of TUG in elderly with knee OA</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Table 2. The test-retest reliability in elderly with knee OA</strong></td>
</tr>
<tr>
<td><strong>Variable</strong></td>
</tr>
<tr>
<td>Cronbach’s alpha</td>
</tr>
<tr>
<td><em>Intraclass correlation</em></td>
</tr>
</tbody>
</table>
Table 2 shows that the test-retest reliability of the TUG test in elderly people with knee OA was stated very high reliability (excellent) with an internal consistency ($\alpha>0.9$). The TUG retest test was carried out at a time interval of 15 minutes resulting in a test retest reliability value of 0.982 (very high reliability).

### Table 3. Correlation test results between test 1 and test 2 for test-retest reliability

<table>
<thead>
<tr>
<th>Variable</th>
<th>Result</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation test between test 1 and 2</td>
<td>$&lt;0.001$</td>
<td>Significant</td>
</tr>
<tr>
<td>Correlation coefficient</td>
<td>$r=0.990$</td>
<td>Very strong</td>
</tr>
</tbody>
</table>

Table 3 shows the analysis of the relationship between tests 1 and 2 using the Spearman Rho test, in which there is a significant correlation between tests 1 and 2 ($p<0.05$) and has a very strong correlation ($r>0.9$).

### Interrater reliability of TUG in elderly with knee OA

Table 4 describes the inter-rater reliability of the TUG test in knee OA patients is stated to have very high reliability (excellent) with an internal consistency alpha was more than 0.9. TUG inter-rater reliability was carried out by 2 raters simultaneously and produced a test-retest reliability value of 0.995 with very high-reliability.

### Table 4. TUG inter-rater reliability test results in knee OA patients

<table>
<thead>
<tr>
<th>Variable</th>
<th>Result</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cronbach's alpha</td>
<td>0.997</td>
<td>Excellent</td>
</tr>
<tr>
<td>Intraclass correlation</td>
<td>0.995</td>
<td>Excellent</td>
</tr>
</tbody>
</table>

Table 5 shows the results of the analysis of the relationship between the measurement results of raters 1 and 2 using the Spearman rho test, which shows that it is significant ($p<0.05$) and has a very strong correlation ($r>0.9$) applied to elderly people with knee OA.

### Table 5. Correlation test between rater 1 and rater 2 of the TUG test

<table>
<thead>
<tr>
<th>Variabel</th>
<th>Result</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation test</td>
<td>$&lt;0.001$</td>
<td>Significant</td>
</tr>
<tr>
<td>Correlation coefficient</td>
<td>$r=0.990$</td>
<td>Very strong</td>
</tr>
</tbody>
</table>

### DISCUSSION

#### Characteristics Respondents

The average age of participants in this study was over 60 years. This follows the explanation (Hsu & Siwiec, 2018) that OA is a degenerative disease that is often found in elderly people. It is estimated that the prevalence of knee OA in people aged 60 years or older is around 10% (men) and 13% (women) (Zhang & Jordan, 2010). One of the most significant risk factors for OA is ageing. Many studies show that radiographic and symptomatic OA of the hands, hips, spine, and knees improve with age (Shane Anderson & Loeser, 2010). Its high prevalence makes OA a major cause of pain and disability in the elderly (CDC, 2010).

This research is dominated by women. Women have a high risk of developing knee OA because hormones and knee anatomy are different from men. The development and persistence of inflammatory cytokines will occur in women's knees. This is a secondary effect of hormonal influence. Decreased postmenopausal estrogen hormones are associated with increased production of inflammatory cytokines such as interleukin-6 (IL-6) (Boyan et al., 2012).

Women have narrower femurs, thinner patellas, larger quadriceps angles, and size differences in the tibial condyle than men. This causes different kinematics, which predisposes the female sex to be more likely to develop OA disease, and ultimately causes a higher prevalence of OA in women (Hame & Alexander, 2013; Long et al., 2020).
According to rheumatology, OA is a disorder characterized by significant loss of cartilage and narrowing of the joint space (Sokolove & Lepus, 2013). Chronic inflammation in OA will damage the articular cartilage and cause various changes in the synovium, joint capsule, synovial fluid, ligaments and tendons (Di Nicola, 2020). As the disease progresses inflammation increases intra-articular pressure within the joint, and significant cartilage damage, allowing synovial fluid to enter the subchondral bone marrow. This causes the production of bone marrow pseudocysts which result in bone erosion and remodelling as well as osteophyte formation (Aigner et al., 2015; Goldring & Goldring, 2010).

Clinically, articular cartilage micro-injuries and cartilage debris in synovial fluid may be present for a long time before damage can be detected by MRI or arthroscopy (Sokolove & Lepus, 2013). To identify individuals who have early indications of joint disease and initiate early therapeutic approaches, Luyten et al. proposed classification criteria for Early Knee Osteoarthritis (EKOA) in 2012 (Luyten et al., 2012; Madry et al., 2016).

Some authors suggest that synovial inflammation may be the primary cause of structural changes. Abnormalities in joint biomechanics and chronic postural defects certainly play a key role in the induction of OA (Di Nicola, 2020). Although not the only tissue affected by OA-associated inflammation, the synovial membrane appears to be a key site of early inflammatory changes (Sellam & Berenbaum, 2010).

Although degeneration and loss of cartilage are key aspects of OA, increased damage to all joint components indicates that the disease affects the joint as an “organ” (Loeser et al., 2012). According to current research, several chronic degenerative diseases with low grades sets the stage for some chronic degenerative diseases. Chronic inflammation in OA is a major factor in the development of joint degeneration (Sokolove & Lepus, 2013).

The main causes of chronic inflammation in OA are joint damage and chronic biomechanical derangement caused by injury (trauma, overuse, hypermobility, postural misalignment, etc.). They call it a ‘vicious cycle’ consisting of local tissue damage, inflammation and failure of tissue repair resulting in chronic synovitis and fibrosis, loss of cartilage and further joint degeneration. Scanzello et al. interestingly linked the occurrence of OA to persistent injuries (Scanzello et al., 2008).

Looking at the research results, OA status occurs more frequently bilaterally. Recent data show that 80% of patients with unilateral knee OA will become bilateral within 12 years (Metcalfe et al., 2012). A previous study from (Spector et al., 1994) with a 2-year follow-up found that 34% of patients with unilateral knee OA would develop the same condition in the opposite knee (with Kellgren & Lawrence OA grade 2). Historically, OA was considered a chronic degenerative disease centered on cartilage damage, and was progressive. Recent research and clinical evidence show that OA gradually affects all joint tissues at different times (Dieppe, 2011; Martel-Pelletier & Pelletier, 2010). This is the basis why the average duration of knee OA sufferers in this study was 3.5 years. This is associated with the fact that OA cannot be cured because it is a degenerative and progressive disease.

Osteoarthritis can get worse over time. Each person's clinical symptoms will be different depending on the severity (Hsu & Siwiec, 2018). The average grade of knee OA sufferers in this study was 2 for the right knee and left knee. The grade of knee OA experienced is linear with the duration and status of the OA. OA usually develops gradually. However, after 4 years, 3.4% of individuals showed radiographic signs of accelerated progression of knee OA (Driban et al., 2020; Driban et al., 2014). Radiographic bone changes or osteophyte formation in OA can be classified using the Kellgren and Lawrence classification to determine the grade of OA that occurs. According to the results of research from (Driban et al., 2020) which was carried out for 15 years with evaluations every 5 years in women with knees without radiographic signs of knee OA (Kellgren & Lawrence grade <2) the results were that Accelerated Knee Osteoarthritis (AKOA) of respondents who experienced accelerated progression of knee OA, grade increased to ≥3 (presence of osteophytes and loss of joint space) within 5 years (Driban et al., 2016). Typical knee OA sufferers experience an increase in radiographic grade within 5 years (KL = 0 to 1, 0
to 2, 1 to 2 and 3). No knee OA had the same grade for more than 5 years.

As time increases, the grade of OA will increase progressively with increasing age, so that the musculoskeletal system will experience degeneration or decline which will result in tissue physiology also decreasing. If good treatment or management is not carried out, it will become a chronic disease which will cause impairment. Complex problems for the patient will reduce the quality of the patient's daily abilities. International agreements also state that it is necessary to start providing OA treatment early in the early stages of the disease when the progression is still reversible (Di Nicola, 2020).

Bilateral knee pain was an independent risk factor for poor physical function according to the results of an analysis with a large longitudinal cohort of individuals at high risk for OA, even when pain intensity was taken into account (White et al., 2010). This means that the level of pain ranging from mild to severe will have an impact on physical abilities. This may be caused by the loss of "good limb" or normal limb function to compensate during functional activities (White et al., 2010).

Patients with unilateral knee OA do not have normal biomechanics in the contralateral limb and will have an asymmetrical gait when experiencing bilateral OA (Creaby et al., 2012; Metcalfe et al., 2013). This happens because of muscle imbalance. Reduced quadriceps activation and greater hamstring activation are the underlying causes of increased co-activity. Altered muscle activation in the knee can disrupt the normal load distribution in the knee and accelerate disease progression (Hortobágyi et al., 2005).

Patient treatment may differ depending on which side is affected, unilateral or bilateral. The study of (Metcalf et al., 2012) highlights that the majority of patients with unilateral disease would benefit from interventions aimed at preventing the possibility of the disease appearing on the other side of the joint or its bilateral occurrence. So it would make sense that both knees of unilateral knee OA sufferers should be managed well even when only one side is bothering them (Creaby et al., 2012).

The average TUG was 16.8 seconds in this observation, where this value is included in the criteria below the average normal dynamic balance ability in elderly OA patients. That, according to (Shumway-Cook et al., 2000) states that elderly people with TUG values below or 13.5 seconds are stated to have good balance and/or have a low risk of falling. However, in this study, the average elderly person was categorized as having a high risk of falling.

**Test-retest reliability**

Test-retest reliability is a repetition of dynamic balance checks using the TUG test carried out by one rater or examiner, where the examiner repeats the TUG test after 15 minutes. In this research, the results showed that internal consistency (Cronbach's Alpha) was 0.991, which means very high internal consistency. The internal consistency of the TUG test in this study means that the TUG test is a test that has a very high construct for measuring dynamic balance in patients with knee OA.

Meanwhile, the ICC value of the TUG test for knee OA sufferers in this study was 0.982, which means the test-retest reliability value is very high (excellent) according to the interpretation of (Koo & Li, 2016). Meanwhile, the correlation value between the results of test 1 and test 2 produces a significance value of less than 0.05, and the correlation coefficient between the results of test 1 and test 2 is 0.962, meaning the correlation is very strong (Cicchetti, 1994). This is following research results from (Alghadir et al., 2015) which stated that the intra-rater reliability of TUG in elderly people with knee OA had a reliability of 0.97. The Osteoarthritis Research Society International (OARSI) also suggests five performance-based tests of body function that can be used, including TUG for people with knee OA (Dobson et al., 2013). With the results of the analysis, it can be said that the TUG test is reliable for measuring dynamic balance abilities in patients with knee OA and has a very high test-retest reliability value.

**Interrater reliability**

Inter-rater is a dynamic balance test using the Timed Up and Go test which is carried out by two raters or examiners at the same time. The purpose of this test is to find out that the consistency of a measuring instrument used by two raters will produce the same results. This is proven by the
Cronbach's alpha value of 0.997 (excellent) in this study. The ICC in this study stated that the measurement results between rater 1 and rater 2 had very high consistency. Or it can be said that the results between rater 1 and rater 2 have an agreement or agreement that is excellent according to the interpretation of (Koo & Li, 2016) with a value of 0.995. This shows that TUG is very consistent when carried out by more than one rater. This is in accordance with research (Alghadir et al., 2015) that the inter rater reliability of TUG in knee OA sufferers produced a value of 0.96 (excellent).

The correlation between the measurement results of rater 1 and rater 2 was significant p-value (<0.001) (Kirkwood & Sterne, 2001) and the correlation coefficient between the results of rater 1 and rater 2 is r=0.990, meaning the correlation is very strong (Cicchetti, 1994) . A significant p-value means that there is evidence against H0. So H0 will be rejected, and Ha accepted. It can be concluded that the TUG is reliable in terms of inter rater for measuring dynamic balance in knee OA patients. A correlation coefficient greater than zero and close to 1 indicates a positive relationship. This means that there is a unidirectional relationship between the two variables. In this case, rater 1 and rater 2 produced almost the same measurements.

**CONCLUSION AND RECOMMENDATION**

**Conclusion**

The Timed Up and Go test is reliable and recommended in terms of test retest and inter rater to measure dynamic balance in elderly people with knee osteoarthritis with excellent results.

**Suggestions**

It is hoped that this research can become reference material for subsequent research. It is recommended to pay attention to other factors that can influence the severity of knee OA which causes dynamic balance disorders such as Body Mass Index and variations in pain scales. Because pain is a person's subjective feeling, it is necessary to limit the level of pain.

**REFERENCES**


Luyten, F. P., Denti, M., Filardo, G., Kon, E., & Engebretsen, L. (2012). Definition and


Surgeons, 20(10), 668-669. https://doi.org/10.5435/JAAOS-20-10-668


https://doi.org/10.1016/j.clinbiomech.2004.08.004


osteoarthritis: is osteoarthritis a chronic wound? *Curr Opin Rheumatol*, 20(5), 565-572. https://doi.org/10.1097/BOR.0b013e32830aba34


