Agreement between Estimated VO₂ max by 6-Minute Walk Test and Non-Exercise Equation in Physiotherapy Students

Rucha Vinay Rayas, Jaimala Vijay Shetye and Amita Anil Mehta

Abstract
Aim: The aim of this study was to find out whether an agreement exists between two methods of estimating VO₂ max. One method, developed by Burr et. Al., (2011) predicted VO₂ max using Six minute walk test (VO₂ max-6MWT), while the other method used a non-exercise equation (VO₂ max-NEx) developed by Jackson et. al.,(1990). Material and Method: One hundred sixty two subjects (N= 162) comprising of 22 male and 140 female physiotherapy students, between ages 18 to 23 years volunteered for the study. Each subject answered the Physical Activity Rating (PA-R) scale and underwent a 6-minute walk test. Collected data was used to estimate VO₂ max. Results: Mean age, BMI, PA-R score and 6-minute walk distance was 20.38 ± 1.3 years, 22.21 ± 4.14 kg/m², 2.95 ± 1.82 and 616.07 ±49.83 meters respectively. Physical activity level was ‘modest’ in 40.12% subjects, ‘low’ in 30.86% and ‘heavy’ in 29.01% subjects. Mean VO₂ max-6MWT and VO₂ max-NEx was 42.98 ± 4.34 ml/kg/min and 38.73 ± 6.07 ml/kg/min respectively. Mean Difference (d) between the two measures was 4.25 ± 4.11 ml/kg/min which exceeded the maximum acceptable difference of 3 ml/kg/min decided a priori. Standard error of mean was 0.32 ml/kg/min. Standard error of limits of agreement was 0.56 ml/kg/min. Bland Altman graphical analysis showed the line of equality (X axis) did not fall within the confidence interval of the mean difference. Conclusion: VO₂ max estimated from 6-minute walk test and VO₂ max estimated from non-exercise equation show no agreement with each other.

Rucha Vinay Rayas
MPT- Cardiovascular & Respiratory Physiotherapy
P. T. School & Centre, Seth GSMMC and KEMH, Parel, Mumbai India
E-mail: rucha.rayas@gmail.com

Jaimala Vijay Shetye
Associate Professor
P. T. School & Centre, Seth GSMMC and KEMH, Parel, Mumbai India
E-mail: jaimalavshetye@gmail.com

Amita Anil Mehta
Professor
P. T. School & Centre, Seth GSMMC and KEMH, Parel, Mumbai India
E-mail: amitaam@gmail.com

Key Words: Aerobic Fitness, Oxygen Consumption, Bland Altman, Physical Activity Rating

DOI: 10.18376/jesp/2019/v15/i1/111319

Introduction
Physical fitness is the ability to perform occupational, recreational, and daily activities without becoming unduly fatigued (Heyward & Gibson, 2014). Cardiopulmonary endurance / Cardiorespiratory fitness (CRF) is one of the ‘health related’ components of physical fitness which refers to the dynamic exercise performance involving large muscles for prolonged periods, which...
depends upon functional status of cardiovascular, respiratory as well as musculoskeletal systems (ACSM, 2010). Hence, determination of CRF gives an idea of overall fitness of an individual. The widely accepted criterion measure of CRF is maximal amount of oxygen that can be utilised by the body during strenuous exercise, termed as VO₂max (ACSM, 2010). The gold standard for measuring VO₂max is a computerized maximal graded exercise stress test (GxT) commonly performed on a treadmill or bicycle ergometer with gas analysis. When the use of such laboratory methods is not feasible, field tests or non-exercise methods can provide information about VO₂max using prediction equations (McArdle et. al., 2010). Multiple studies have demonstrated majority physical therapy students to have good to average or poor aerobic fitness levels. Balogun et. al., (1987) assessed the aerobic capacity of 50 female physical therapy students with a submaximal graded bicycle ergometer test. The estimated VO₂max values were in the ‘good’ (38-48 ml/kg/min) to ‘average’ (31-37 ml/kg/min) range of CRF based on classification mentioned in ‘The Physical Fitness Specialist Certification Manual’ (1997). More recent studies on physiotherapy students from Indian population which have used the Harvard step test as a measure of CRF provide similar findings. A study on 250 students from Punjab and Haryana (Multani, et al., 2013) found 66.8% subjects to have ‘poor’ levels of CRF. Similarly, 92% of a sample of 50 students from School of Physiotherapy, R. K. University (Bhansali & Bharmal,2015) were classified as having ‘poor’ fitness levels. Few studies have measured energy expenditure of physical therapists during their workplace activities in various sub-fields/ specialty areas. Balogun (1986) and colleagues attempted to determine the energy cost of physical therapists during clinical practice using an accelerometer. Energy expenditure was found to be 2.4 METs which translates to ‘light’ levels. However, a commentary on this article quotes that these results “seriously underestimate the energy output required of the physical therapist” which may be attributed to small study sample, accelerometer limitations and nature of work in clinical practice. Although there is no specific objective measure of job demand of physical therapy profession currently available, it is largely agreed upon that physical therapists are health professionals requiring good/high levels of physical fitness, adequate enough to meet their job demands (Balogun, 1987; Angell, et al., 1999; Multani, et al., 2013; Bello, et al., 2016). Also, the academic duties of physiotherapy students require prolonged activities during practical learning sessions and clinical postings. Considering that the PT profession has high work demands, it is necessary to understand the current fitness levels in PT students who will become future professionals in this field. Understanding of their own fitness levels, early in professional training course may help encourage physiotherapy students to participate in various forms of exercise. There is a need to conduct surveys to understand present level of fitness in this population. In large scale epidemiological studies, it may be necessary to document cardio-respiratory fitness of an entire community, such as a community of physiotherapy students. In such circumstances, it is impractical to perform a treadmill test or a field test on each and every individual. The need for the present study is to explore the applicability of Physical Activity Rating (PA-R) based non-exercise VO₂ max prediction equation developed by Jackson et. al., (1990) by determining its degree of agreement with VO₂ max estimated using 6-minute walk test (Burr et. al., 2011) in physiotherapy students for future use in fitness surveys, the results of which may allow extending the study to patient populations.

**Material and Methods**

The study was conducted on 162 students of Bachelor of Physiotherapy course at Physiotherapy School and Centre, Seth GSMD and KEM Hospital, Mumbai and their age ranged 18 to 23 years. Total 162 subjects participating in the study comprised of 140 females and 22 males. After obtaining written informed consent, each participant was screened using Physical Activity Readiness Questionnaire (PAR-Q). Fulfillment of inclusion and exclusion criteria was determined. Each subject’s demographic and anthropometric data (height, weight, BMI) was noted. The self-
report Physical Activity Rating Scale (PA-R) (Jackson, et al., 1990) was read and filled in by each subject in a comfortable seated position. Any queries were resolved by investigators when needed.

**6-minute walk test (6MWT) procedure**

6MWT as per American Thoracic Society (ATS) guidelines (2002) was performed indoors along a 30m long corridor with ends of the walkway marked with 2 cones and a floor marking (red tape) for the starting point. Standard instructions and encouragement were provided before and during the test respectively. An early afternoon test was done at least 1-hour post meal. Practice test was given at around same time of the day for each subject respectively, to decrease variability and allow for learning effect of 1 practice test. Final test was conducted within 1-week of practice test. Subjects wore lose comfortable clothing with appropriate footwear and refrained from caffeine consumption or performance of exercise at least 2 hours prior to the test. Vital parameters, which included respiratory rate (RR), resting heart rate (RHR), systolic and diastolic blood pressure and rate of perceived exertion (RPE) were documented at rest, immediately after the test and every 2 minutes thereafter till parameters returned to pre-test level. The information entered in the case record forms and PA-R score were used to calculate predicted VO₂ max (VO₂ max-6MWT and VO₂ max-NEx). Data was entered and organised using Microsoft Excel office 365 and analysed using SPSS 16.0 (SPSS Inc, Chicago). Level of significance was set at p <0.05 at confidence interval of 95%. Mean and standard deviation was calculated for age, BMI, PA-R score, 6-minute walk distance. Bland Altman analysis (Altman & Bland, 1983) was used to determine the agreement between VO₂ max-NEx and VO₂ max-6MWT and data was further analysed using graphical approach (B & A Plot).

**Results**

Total 162 subjects participating in the study comprised of 140 females and 22 males with mean age of 20.38 ± 1.3 years. Mean BMI, PA-R score and 6-minute walk distance was 22.21 ± 4.14 kg/m², 2.95 ± 1.82 and 616.07 ±49.83 meters respectively (Table 1).

### Table 1. Descriptive Statistics (N=162)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Units</th>
<th>Mean ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Years</td>
<td>20.38 ± 1.3</td>
</tr>
<tr>
<td>BMI</td>
<td>kg/m²</td>
<td>22.21 ± 4.14</td>
</tr>
<tr>
<td>PA-R score</td>
<td>-</td>
<td>2.95 ± 1.82</td>
</tr>
<tr>
<td>6 MWD</td>
<td>Meters</td>
<td>616.07 ± 49.83</td>
</tr>
</tbody>
</table>

The data did not follow normal distribution. Most subjects had ‘modest’ physical activity level (40.12%) followed by ‘low’ physical activity level (30.86%). Remaining subjects had ‘heavy’ physical activity levels (29.01%). Mean VO₂ max-6MWT and VO₂ max-NEx was 42.98 ± 4.34 ml/kg/min and 38.73 ± 6.07 ml/kg/min respectively (Table 2). Mean Difference (Δ) between VO₂ max-6MWT and VO₂ max-NEx was 4.25 ± 4.11 ml/kg/min which exceeded the maximum acceptable difference of 3 ml/kg/min decided a priori. Standard error of mean was 0.32 ml/kg/min. Standard error of limits of agreement was 0.56 ml/kg/min. Figure 1. Bland Altman graphical analysis (B & A plot) showed the line of equality (X-axis) did not fall within the confidence interval of the mean difference.
Table 2. Mean ± standard deviation of VO\textsubscript{2}max predicted from 6MWT and Non-exercise equation

<table>
<thead>
<tr>
<th>Variable</th>
<th>Females (140)</th>
<th>Males (22)</th>
<th>Total (N=162)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VO\textsubscript{2}max-6MWT</td>
<td>42.09 ± 3.61</td>
<td>48.40 ± 4.72</td>
<td>42.98 ± 4.34</td>
</tr>
<tr>
<td>VO\textsubscript{2}max-NEx</td>
<td>37.18 ± 4.35</td>
<td>48.19 ± 6.79</td>
<td>38.73 ± 6.07</td>
</tr>
</tbody>
</table>

Figure 1. Bland Altman Plot for total sample (N=162) shows a horizontal solid black line = Mean Difference (\(\bar{d}\)) = 4.25 ml/kg/min.

In Figure 1, the dotted blue lines above and below the solid black line are upper and lower 95% confidence intervals of \(\bar{d}\) at 4.89 ml/kg/min and at 3.61 ml/kg/min respectively. The dotted black line (above mean difference) denotes upper limit of agreement (LOA) which is 14.86 ml/kg/min. This line is bound by two dotted blue lines which are its upper and lower 95% confidence intervals at 15.97 ml/kg/min and 13.75 ml/kg/min respectively. Similarly, the dotted black line (below mean difference) denotes lower limit of agreement (LOA) which is -6.37 ml/kg/min. This line is bound by two dotted blue lines which are its upper and lower 95% confidence intervals at -5.26 ml/kg/min to -7.48 ml/kg/min respectively.
In Figure 2, the dotted blue lines above and below the solid black line are upper and lower 95% confidence intervals of $\bar{d}$ at 5.53 ml/kg/min and at 4.31 ml/kg/min respectively. The dotted black line (above mean difference) denotes upper limit of agreement (LOA) which is 14.37 ml/kg/min. This line is bound by two dotted blue lines which are its upper and lower 95% confidence intervals at 15.44 ml/kg/min and 13.30 ml/kg/min respectively. Similarly, the dotted black line (below mean difference) denotes lower limit of agreement (LOA) which is -4.54 ml/kg/min. This line is bound by two dotted blue lines which are its upper and lower 95% confidence intervals at -3.47 ml/kg/min to -5.61 ml/kg/min respectively.

Figure 3. Bland Altman Plot for Males (n=22) shows a Horizontal solid black line = Mean Difference ($\bar{d}$) = 0.21 ml/kg/min.
In Figure 3, the dotted blue lines above and below the solid black line are upper and lower 95% confidence intervals of \( \bar{d} \) at 2.14 ml/kg/min and at -1.72 ml/kg/min respectively. The dotted black line (above mean difference) denotes upper limit of agreement (LOA) which is 11.44 ml/kg/min. This line is bound by two dotted blue lines which are its upper and lower 95% confidence intervals at 14.79 ml/kg/min and 8.09 ml/kg/min respectively. Similarly, the dotted black line (below mean difference) denotes lower limit of agreement (LOA) which is -11.02 ml/kg/min. This line is bound by two dotted blue lines which are its upper and lower 95% confidence intervals at -7.67 ml/kg/min to -14.37 ml/kg/min respectively.

### Table 3. Bland Altman analysis parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Standard Error formula</th>
<th>Standard error (SE)</th>
<th>t value (t)</th>
<th>Confidence (SE*t)</th>
<th>Confidence Intervals (CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean difference (( \bar{d} ))</td>
<td>4.25</td>
<td>( \sqrt{\frac{s^2}{n}} )</td>
<td>0.32</td>
<td>1.98</td>
<td>0.63</td>
<td>4.88 - 3.62</td>
</tr>
<tr>
<td>Upper limit of agreement (( \bar{d} + 2.58s ))</td>
<td>14.86</td>
<td>( \sqrt{\frac{s^2}{n}} )</td>
<td>0.56</td>
<td>1.98</td>
<td>1.11</td>
<td>15.97 - 13.75</td>
</tr>
<tr>
<td>Lower limit of agreement (( \bar{d} - 2.58s ))</td>
<td>-6.37</td>
<td>( \sqrt{\frac{s^2}{n}} )</td>
<td>0.56</td>
<td>1.98</td>
<td>1.11</td>
<td>-5.26 - -7.48</td>
</tr>
</tbody>
</table>

n = 162, degrees of freedom (df) = 161, standard deviation (s) = 4.11 ml/kg/min

### Table 4. Bland Altman Analysis parameters for females (n=140)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Standard Error formula</th>
<th>Standard error (SE)</th>
<th>t value (t)</th>
<th>Confidence (SE*t)</th>
<th>Confidence Intervals (CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean difference (( \bar{d} ))</td>
<td>4.92</td>
<td>( \sqrt{\frac{s^2}{n}} )</td>
<td>0.31</td>
<td>1.98</td>
<td>0.61</td>
<td>5.53 - 4.3</td>
</tr>
<tr>
<td>Upper limit of agreement (( \bar{d} + 2.58s ))</td>
<td>14.37</td>
<td>( \sqrt{\frac{s^2}{n}} )</td>
<td>0.54</td>
<td>1.98</td>
<td>1.07</td>
<td>15.44 - 13.38</td>
</tr>
<tr>
<td>Lower limit of agreement (( \bar{d} - 2.58s ))</td>
<td>-4.54</td>
<td>( \sqrt{\frac{s^2}{n}} )</td>
<td>0.54</td>
<td>1.98</td>
<td>1.07</td>
<td>-3.47 - 5.6</td>
</tr>
</tbody>
</table>

n= 140, degrees of freedom (df)= 139, standard deviation (s) = 3.67 ml/kg/min
Table 5. Bland Altman Analysis parameters (Males n=22)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Standard Error formula</th>
<th>Standard error (SE)</th>
<th>t value (t)</th>
<th>Confidence (SE*4)</th>
<th>Confidence Intervals (CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean difference ((\bar{d}))</td>
<td>0.21</td>
<td>(\sqrt{s^2/n})</td>
<td>0.93</td>
<td>2.08</td>
<td>1.93</td>
<td>2.14, -1.72</td>
</tr>
<tr>
<td>Upper limit of agreement ((\bar{d} + 2.58s))</td>
<td>11.44</td>
<td>(\sqrt{s^2/n})</td>
<td>1.61</td>
<td>2.08</td>
<td>3.35</td>
<td>14.79, 8.09</td>
</tr>
<tr>
<td>Lower limit of agreement ((\bar{d} - 2.58s))</td>
<td>-11.02</td>
<td>(\sqrt{s^2/n})</td>
<td>1.61</td>
<td>2.08</td>
<td>3.35</td>
<td>-7.67, 14.37</td>
</tr>
</tbody>
</table>

\(n=22\), degrees of freedom (df) = 21, standard deviation (s) = 4.35 ml/kg/min

Table 6: Physical Activity Rating (PA-R) Score Categorization

<table>
<thead>
<tr>
<th>PA-R score</th>
<th>No. of subjects out of total 162</th>
<th>Percentage of subjects (%)</th>
<th>Activity level</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 to 1</td>
<td>50</td>
<td>30.86</td>
<td>Low</td>
</tr>
<tr>
<td>2 to 3</td>
<td>65</td>
<td>40.12</td>
<td>Modest</td>
</tr>
<tr>
<td>4 to 7</td>
<td>47</td>
<td>29.01</td>
<td>Heavy</td>
</tr>
</tbody>
</table>

Discussion

The aim of this study was to find out whether an agreement exists between two methods of estimating VO\(_2\)max. One method predicted VO\(_2\)max using an exercise test i.e. Six minute walk test (VO\(_2\)max-6MWT) using an equation developed by Burr et al., (2011) while the other method used a non-exercise equation (VO\(_2\)max-NEx) developed by Jackson et al., (1990). A total of 162 physiotherapy students participated in the study and comprised of 22 males (14%) and 140 females (86%) between the ages of 18 to 23 years. Analysis of data revealed that the mean difference (\(\bar{d}\)) between VO\(_2\)max-6MWT and VO\(_2\)max-NEx was 4.25 ml/kg/min ± 4.11 ml/kg/min. This indicates that on an average, the 6MWT based method estimates VO\(_2\)max 4.25 ml/kg/min more than non-exercise method in the studied sample. This difference was more than the maximal allowable difference of 3 ml/kg/min decided a priori. In Figure 1, since the line of equality (X-axis) did not fall within the confidence interval of the mean difference, it can be concluded that, the VO\(_2\)max estimated by 6 MWT based prediction method and the Non-exercise prediction method are constantly different from each other. When an indirect measure of VO\(_2\)max such as a field test is assessed for agreement with a maximal GxT (gold standard), the prediction error / standard error estimate (SEE) is approximately < 5.0 ml/kg/min (McArdle, et al., 2010). In the present study, SEE was found to be 5.914 ml/kg/min which exceed the allowable error. Hence, it was concluded that agreement did not exist between VO\(_2\)max-6MWT and VO\(_2\)max-NEx for the present study sample.
Agreement analysis was then done for both genders separately. Details of Bland Altman analysis for females (n=140) with B & A graph for the same is shown in Table 4 and Figure 2 respectively. Similar to findings of total study sample, females also showed no agreement between VO₂max-6MWT and VO₂max-NEx. The SEE for females was 6.11 ml/kg/min was exceeded the allowable error of 5 ml/kg/min as mentioned before. It was interesting to note that analysis done separately for males (n=22) showed that VO₂max-6MWT and VO₂max-NEx demonstrated agreement with each other. Details of Bland Altman analysis for males with B & A graph of the same is shown in Table 5 and Figure 3 respectively. The mean difference between the two methods was only 0.21 ml/kg/min and as seen in Figure 3, the line of equality (X-axis) lies within the confidence interval of the mean difference. Also, SEE for males was 4.46 ml/kg/min which was less than the acceptable prediction error. Thus, an agreement exists for the group of 22 males studied. The clinical significance of this finding however is questionable since the gender distribution was unequal and sample of males was small (n=22). Further study is warranted. The non-exercise prediction equation in the study by Jackson et. al., (1990) showed strong correlation of VO₂max-NEx with measured VO₂ max. This study comprised of a large study sample with around 90% male and 10% female participants. This may help explain why an agreement between VO₂max-6MWT and VO₂max-NEx was demonstrated in males in our study. This needs to be further studied. It is known that 6-minute walk distance (6MWD) correlates positively with actual VO₂max obtained from GxT in various patient populations (Eaton et. al., 2005; Ross et. al., 2010; Harmsen et. al., 2017). The correlation between these two variables is inconsistent in healthy population (Ming Ma & Lin Ma, 2009; Andersson & Nilsson, 2011; Burr, et. al., 2011; Nusdwinuringtyas et. al., 2011; Zhang et. al., 2017). The present study used estimated VO₂max and determined its correlation with 6MWD. It was found that 6MWD had weak positive correlation with VO₂max-NEx with a Pearson correlation coefficient of 0.407 (p < 0.001). Studies showing significant positive correlations between 6MWD and VO₂max by GxT such as those by Burr et. al.,(2011), Zhang et al(2017) and Nusdwinuringtyas et. al., (2011), have samples with middle to old aged population. Whereas studies by Andersson and Nilsson (2011) or Ming Ma and Lin Ma(2009) have shown very weak positive and no correlations respectively in younger subjects. This may explain the present study findings of weak correlation, since the subjects were young college students. Whether these differences are merely due to different ages is difficult to comment upon. These inconsistencies may be due to non-adherence to specific ATS guidelines (2002) while performing 6MWT by various researchers which alters 6MWD in different study settings. The methods in some above mentioned studies were different from those recommended by ATS guidelines. Burr et. al., (2011) did not provide any encouragement to study participants while they performed the 6MWT. Also, Nusdwinuringtyas (2011) used a 15m track instead of 30m during 6MWT. Such differences in test procedure can affect outcomes giving rise to error in result interpretation. PA-R score of study participants, which indicates the self-reported physical activity level in the past one month, was expected to correlate well with 6MWD however, it was seen that these two variables showed a very weak positive correlation (Spearman ρ = 0.184, p < 0.05). Studies in the past have shown self-reported physical activity levels to correlate well with 6MWD. These studies have used questionnaires such as the International Physical Activity Questionnaire (IPAQ) (Gurses et. al., 2018) or have obtained information about habitual activity levels through a structured interview (Lord & Menz, 2002). These are more elaborate methods of assessing self-reported physical activity levels. The Physical Activity Rating (PA-R) scale used in the present study is an 8-point Likert scale. A single score could depict a wide range of physical activity levels. Hence, the scale scores may not have been able to adequately reflect functional capacity in terms of 6MWD. Secondly, since it is a self-report scale the understanding with which subjects responded cannot be controlled. Use of a more elaborate physical activity questionnaire is advocated while conducting future research on this.
topic. Mean score of Physical Activity Rating (PA-R) (0 to 7) was 2.95. Most subjects (40.12%) reported ‘modest’ activity levels followed by ‘low’ activity levels (30.86%). This finding is similar to those of previous studies on physiotherapy student population that have used IPAQ to measure self-report physical activity. Some recent studies by Ranasinghe et. al., (2016) and Dabholkar et. al., (2018) have shown physiotherapy students to have moderate to poor physical activity levels/habits. The present study results add to this, further highlighting the need to assess and improve physical fitness of physiotherapy students on a larger scale.

Conclusion
VO2max estimated from 6-minute walk test and VO2max estimated from non-exercise equation show no agreement with each other.

Acknowledgement
I wish to express my sincere thanks to all participants of this study, departmental staff, my colleagues, statistician and friends for their valuable guidance and encouragement.

References


Conflict of Interest: None declared