Neuromechanical Justification of Lower-Limb Functional Tests for a Return to Running: A Muscle Coordination Analysis

Hiroki Saito 1,2*, Ayu Yamano1, Nanae Suzuki1, Kazuya Matsushita 3, Hikaru Yokoyama4, Atsushi Sasaki 5, 6, Tatsuya Takahashi MD 7, Kimitaka Nakazawa PhD 2

1 Department of Physical Therapy, Tokyo University of Technology, Tokyo, Japan
2 Graduate School of Arts and Sciences, Department of Life Sciences, The University of Tokyo, Tokyo, Japan
3 Aobadai Takeda Orthopedic, Tokyo, Japan
4 Institute of Engineering, Tokyo University of Agriculture and Technology, Tokyo, Japan
5 Graduate School of Engineering Science, Department of Mechanical Science and Bioengineering, Osaka University, Osaka, Japan
6 Japan Society for the Promotion of Science, Tokyo, Japan
7 Toyota Memorial Hospital, Aichi, Japan

Correspondence Author: Hiroki Saito
Department of Physical Therapy, Tokyo University of Technology, Tokyo, Japan

Email: saitohrk@stf.teu.ac.jp
ABSTRACT

Introduction

This study aimed to explore the shared muscle synergies between running and functional tests that are commonly used when considering the return to running (RTR) after sports-related injuries. We hypothesized that shared muscle synergies would differ among tasks, providing insights into prioritizing functional tests in the context of RTR.

Methods

Ten healthy male participants were recruited to perform running and 9 functional tasks and their 16 lower limb and trunk muscle activities were recorded using electromyography (EMG). Non-negative matrix factorization (NMF) was applied to the collected EMG data to explore shared muscle synergies between running and the functional tasks. We compared the percentages of shared synergies and temporal patterns between running and each functional test.

Results

Although all functional tests exhibited shared muscle synergies with running, the walk (75% [40%-100%]), single leg hops with 30% of maximum distance (SLH30) (60% [20%-100%]), and stepup (63% [0%-100%]) tasks displayed significantly higher percentages of shared synergies compared to other tests. However, significant
differences in temporal patterns were observed between running and all functional 
tasks, indicating varying activation profiles of shared muscle synergies.

Conclusion

Although all functional tests shared muscle synergies with running, variations in the 
degree of shared synergies and temporal patterns imply that walking, SLH30, and step-
up tests may be the most beneficial in predicting running behavior post-ACL injuries.

However, functional tests cannot fully replicate running dynamics, suggesting the need 
for a careful interpretation when assessing readiness for RTR.

keyword: Return to run, functional tests, muscle synergies.
Returning to running (RTR) represents a key milestone in the rehabilitation process following sports-related injuries (1-4). The time frame for RTR can vary greatly; for instance, in the case of anterior cruciate ligament (ACL) injuries, the median time frame for RTR is approximately 12 weeks postoperatively, with a range of 5-39 weeks (1, 2). This phase signifies the transition from early and mid-stage rehabilitation—focused on restoring basic knee function such as regaining adequate knee joint range of motion and resuming everyday activities—to the later stages of rehabilitation. These later stages involve high-intensity training, encompassing activities like jumping, cutting, and sport-specific tasks (5, 6).

Functional tests, which are frequently used as assessment-based criteria, are designed to replicate the physical demands of running. These tests include walk analysis, single leg hops, single leg squats (SLS), and various balance tasks (1, 2). However, it is critical to validate these functional tests as practical benchmarks for facilitating a safe and effective return to running post-injury. This is particularly important for validating tests that share motor control characteristics between running and functional tests. The role of muscles in motor function is interconnected (7, 8). Essentially, motor control of lower limbs entails the coordinated activation of several lower limb
and trunk muscles in response to specific demands requiring both stability and movement. (9). The central nervous system (CNS) is thought to employ an efficient strategy to select the control signal from a large subspace. This is achieved by using a limited set of motor modules or muscle synergies, formed by the flexible combination of muscle activation (10). Thus, exploring the shared muscle synergies between running and commonly used functional tests in clinical settings may provide the neuromechanical rationale for functional tests as RTR criteria following sports-related injuries.

The objective of this study was to determine whether functional tests share muscle synergies with running tasks. Our hypothesis posited that running and functional tests would display shared muscle synergies, but the extent of these synergies would vary between tasks. Additionally, differences in the temporal profiles of these synergies were expected. Such variability could provide critical insights for prioritizing functional tests when evaluating readiness to RTR.
METHODS

Participants

Ten healthy males with a mean (SD) age of 21 (±0.3) years old were recruited from the local university. Each participant provided written informed consent for participation in the study. The study was conducted following the principles of the Declaration of Helsinki and approved by the local ethics committee of the University of Tokyo (746).
Experimental procedures

Participants were asked to freely perform running and nine functional tasks commonly used when considering the return to running following ACL injuries (1), as described in Figure 1. Of note, the average velocities for running and walk were observed to be $2.1 \pm 0.28$ m/s and $1.2 \pm 0.12$ m/s, respectively. Each functional test was repeated five times and the order of the tasks was randomly assigned.

Data collection

Unilateral surface EMG data were recorded from 16 lower limb and trunk muscle groups: rectus abdominis (RA) (3cm lateral to umbilicus)(11), oblique externus (OE) (15cm lateral to umbilicus)(12), erector spinae at L1 (ESL1) (3cm lateral to the L1 spinous process)(11), gluteus maximus (GM), gluteus medius (Gmed), biceps femoris (long head, BF), semitendinosus (ST), tensor fasciae latae (TFL), adductor longus (ADD), rectus femoris (RF), vastus medialis (VM), vastus lateralis (VL), tibialis anterior (TA), gastrocnemius medialis (MG), soleus (SOL) and peroneus longus (POL).

The EMG sensor placements in lower limb were based on SENIAM (surface EMG for a non-invasive assessment of muscles) (13). A wireless EMG system (Trigno Wireless System; DELSYS, Boston, MA, USA) was used to record EMG activity. Each electrode had an inter-electrode spacing of 10 mm. The EMG signals were band-pass filtered (20–
EMG processing

Raw EMG signals were high-pass filtered at 30 Hz to remove motion artifacts, and then demeaned. The signals were then full-wave-rectified and low-pass filtered at 10 Hz, using a fourth-order Butterworth filter. The smoothed EMG envelopes were time-interpolated to generate 200 time points between the start and end points for each
trial so that the EMG data of each trial contributed to the extracted muscle synergies equally.

We created single-task EMG matrices for each task for each participant (that is, the matrix was composed of 16 muscles × 1000 timepoints (the no. of repetitions/cycles (5) × 200 samples)) to extract the muscle synergies for each task. Each EMG from each muscle was normalized to the maximum amplitude across all tasks.

**Independent muscle synergy extraction**

To extract muscle synergies, NMF was applied to the single-task EMG matrix. NMF has previously been described as a linear decomposition technique (14, 15) according to equation (1):

$$M = WC + e$$ (1)

Where $M$ ($m \times t$ matrix, where $m$ is the number of muscles, $t$ is the number of samples [i.e., spatiotemporal profiles of muscle activity]) is a linear combination of muscle synergies, $W$ ($m \times n$ matrix, where $n$ is the number of muscle synergies), $C$ ($n \times t$ matrix, representing temporal patterns), and $e$ is the residual error matrix. We
applied NMF to extract possible \( n \) values from 1 to 16 for each dataset. To estimate the optimal number of muscle synergies, the variance accounted for (VAF) by the reconstructed EMG \((M)\) was calculated at each iteration (16). VAF was defined as 100 \( \times \) the square of the uncentered Pearson's correlation coefficient (16, 17). Considering the local minima inherent in NMF, each synergy extraction was repeated 50 times, and the VAF was calculated at each iteration. Iterations with the highest VAF were maintained (18-21). VAFs > 0.9 were used to identify the optimal number of synergies commonly used in the literature (20-25).

**Shared and specific muscle synergy extraction**

To extract the number of shared and running-specific, test-specific muscle synergies, we used a modified version of the NMF algorithm based on the previous studies (24, 26-28) that simultaneously extracts motor modules that are shared across running and each task and those that are specific to each task from a data matrix containing EMG from both conditions. We defined shared synergies as one that activated in both task and thus, temporal pattern components have non-zero coefficients in both tasks. To identify task-specific muscle synergies, the coefficients, \( C \), corresponding to running are set to zero (i.e., test-specific synergies), and to each task
are set to zero (i.e., run-specific synergies). Detailed descriptions of this method are in elsewhere (24, 26-28). Briefly, as independent synergy extraction, the number of shared, running specific and task-specific synergies for each participant was determined by the minimum number of total muscle synergies that were required so that the VAF exceeded 90%. We defined the percentage of shared muscle synergies between running and each task as the ratio of the number of shared modules over the number of total motor modules across the two tasks. This modification is thought to improve the accuracy shared and task-specific muscle synergies by minimizing the possibility in which the numbers of synergies are underestimated when synergies are extracted from the EMG of each task independently, and compare the similarity of synergies between tasks to identify shared and task-specific synergies (26, 29).

We identified representative shared, running-specific and test-specific synergies across participants using hierarchical clustering analysis (Ward's method, Euclidian distance) of muscle weighting components (19). To determine the optimal number of clusters, we computed the gap statistic (30), which measures the compactness of the clustering achieved against those in reference data sets without any obvious clustering similar to a previous study (31). Reference data sets (N = 500) were initially generated by sampling uniformly from within the bounds of the original muscle-synergy
set; each of them was then clustered by the hierarchical cluster, at 2–20 clusters. The optimal number of clusters was then the smallest number, h, such that

\[ \text{Gap}(h) \geq \text{Gap}(h+1) - \text{sd}(h+1) \]

where \( \text{Gap}(k) \) represents the gap statistic at \( h \) clusters, and \( \text{sd}(h) \) signifies the standard deviation of the clustering compactness within the reference data sets (30).

We defined shared, running-specific and test-specific synergy clusters as having each three types of synergies from \( \geq \frac{1}{2} \) of synergies within a cluster. If the cluster was not contributed by any types of synergies, we defined it as “none”.

**Statistics**

We compared the percentage of shared muscle synergies between tasks. The values were compared using the Friedman test, which is a non-parametric method for multiple comparisons of independent samples, as a normal distribution was not observed in the data (tested using the Shapiro–Wilk test). When the Friedman test
showed significant effects, multiple comparison post-hoc analyses were performed using the Wilcoxon signed-rank test.

Temporal pattern components were compared between running and functional tests using statistical parametric mapping (SPM) (spm1d v0.4.7 for MATLAB, Institute of Neurology, London, UK) (32). Since each single-task EMG matrix contained 5 repetitions/cycles, the extracted temporal pattern components of running and functional tests were converted into an averaged repetition/cycles for each participant before comparing temporal pattern components using SPM analysis.

The p values obtained from were corrected using the false discovery rate (FDR) correction for multiple comparisons (33). The significance level for all tests was set at $p < 0.05$. When there was a significant difference between the groups, effect sizes (ES) were calculated using Cohen’s d (34). We recruited ten participants without an a priori power analysis, thus, we instead conducted a sensitivity analysis in G*Power, which indicated that an effect size of 0.71 would be necessary to obtain a power of 80% at an $\alpha$ of 0.05.
RESULTS

Muscles synergies for running and functional tests

Table 1 presents the VAF values between the groups for running and functional tests. Figure 2 depicts the percentage of muscle synergies shared between running and each respective functional test. The median percentage of shared synergies for each test were as follows: walk (75% [40%-100%]), SLH30 (60% [20%-100%]), SLH60 (50% [28%-75%]), SLH100 (50% [28%-80%]), SLS (32% [16%-50%]), Ybalance (40% [25%-80%]), Hraise (22% [20%-60%]), stepup (63% [0%-100%]), and Tbalance (40% [20%-80%]). The percentage of shared synergies was significantly different among the nine representative speeds in non-runners and runners ($p = 0.000079$ for both groups, Friedman test one-way ANOVA). Statistically significant differences were observed between the following pairings: Walk and SLS ($p = 0.0038$, ES = 2.18), Walk and Hraise ($p = 0.0098$, ES = 2.05), SLH30 and SLS ($p = 0.0017$, ES = 1.56), SLH30 and Hraise ($p = 0.0039$, ES = 1.47), SLH and stepup ($p = 0.0072$, ES = 1.31), and Hraise and stepup ($p = 0.0176$, ES = 1.25).

Figure 3 shows representative shared, running-specific, and test-specific muscle synergies identified by cluster analysis (muscle weighting components (W) and temporal pattern components (C)). All functional tests shared the muscle synergies with
running although the number of shared synergies differed between functional tests. We also identified the running-specific and test-specific muscle synergies in all functional tests except walk task. The SPM analysis found significant differences in temporal pattern components between running and each functional task ($p < 0.05$).
In this study, we applied the NMF algorithm to large-scale and high-dimensional EMG data to investigate shared muscle synergies between running and functional tasks that are commonly used to determine when individuals with ACL injuries can return to running activities. Overall, although all functional tests shared muscle synergies with running, our results suggest that walking, SLH, and step-up tests could be the most beneficial because the percentages of shared synergies between running and these tests were significantly higher than those with other functional tests. However, despite these shared muscle synergies, there were notable differences in the temporal patterns between running and functional tasks. These discrepancies suggest the need for caution when using functional tests alone to predict running capabilities post-injury.

We hypothesized that a high degree of shared synergies between running and functional tests would signify these tests as strong predictors of running behavior, due to similar motor control characteristics, including the coordination of multiple muscles. Firstly, our results showed that the percentage of shared synergies (muscle weighting components) between running and walking was 75% [40%-100%], with all synergies between the two tasks being shared. This aligns with a previous study that suggested consistent muscle weighting components for walking and running, facilitating the
transition from walking to running. (35, 36). Similarly, we found that the step-up test
shared a significant amount of muscle synergies with running, which is reasonable
given that both activities share the same mechanical goal of propelling the body's center
of mass forward. (37-39). This action requires dynamic balance with appropriate
activations of lower and trunk muscles. The SLH test, particularly the SLH30 variant,
exhibited a higher degree of shared muscle synergies with running compared to other
non-jumping tests. Conversely, SLH60 and SLH100 demonstrated a relatively low
number of shared synergies, as they also generated test-specific synergies beneficial for
longer forward jumping.

NMF also extracted temporal pattern components, signifying the activation
profiles of muscle weighting components. Notably, there were significant differences in
some temporal pattern components of muscle synergies between running and all
functional tests. From a neuromechanical perspective, these temporal patterns reflect the
specific timing and intensity of muscle activation, illustrating that even though muscle
synergies may be similar, the way they are employed in different activities can differ
significantly. This discrepancy could be attributed to the unique demands of each task.
Even though similar muscle groups may be engaged (thus the shared synergies), the
coordination, timing, and intensity of muscle activation might not perfectly align. (40,
From a clinical perspective, these differences suggest that functional tests cannot fully replicate running behavior. Specifically, the characteristics of temporal components of walking, SLH30, and stepup tasks that exhibited higher percentages of shared synergies, differed significantly compared to those in running. These disparities in temporal patterns, despite shared muscle synergies, can limit the interpretation of the results, suggesting that the differing activation profiles of shared muscle synergies in these functional tests may impede their predictability of safe running performance.

Clinicians should consider that while these functional tests can help assess readiness for running by reflecting certain shared muscle activation patterns, they do not perfectly mimic the exact dynamics of running. Therefore, clinicians should utilize functional tests in combination with running evaluations in a clinical setting using a treadmill whenever possible. Additionally, other assessments such as knee range of motion, strength, and psychological readiness for running should be taken into account. A comprehensive assessment will aid in detecting basic function for running and deviations from normal running patterns before athletes resume running outside a clinical setting (1, 2).

Steady progress through high-quality rehabilitation is essential for functional recovery (42). Resuming sports activities prematurely can elevate the risk of secondary
injuries (43), while excessively slow progress might adversely affect motivation and psychological preparedness for sports performances (44). Consequently, determining the appropriate timing for a return to running (RTR) is a crucial milestone in the effective rehabilitation continuum for a return to sports (1, 2). Our findings provide a novel rationale for using functional tests in decision-making for RTR. Further exploration into the degree of shared synergies between running and functional tests among individuals with and without lower-limb sports injuries could yield intriguing results. This, along with investigating the relationship between the degree of shared synergies and future injuries, could potentially offer valuable biomarkers for injury prevention.

Note of caution in interpreting the study findings is warranted. Although numerous functional tests mimic the motor control patterns or muscle synergies of running, the running speed employed in this study was relatively low, averaging 2.1 ± 0.28 m/s. This raises uncertainty regarding whether the functional tests used also capture the muscle synergy variations of high-speed running, direction changes, and cutting movements, all of which are high-risk activities for lower limb sports injuries. Indeed, a previous study highlighted distinct muscle synergies at different speeds (19),
If the tests used don’t replicate these running variations, we might not accurately predict running ability and may need to consider other functional test batteries.

CONCLUSION

Our study suggests that the walk, SLH, and step-up tests can be reliable indicators of running behavior due to their shared muscle synergies with running. Despite different temporal pattern components in these tests compared to running, they offer a practical means to assess running ability. However, clinicians should be aware that these functional tests may not fully emulate the physical demands of running. Therefore, these tests should be incorporated as key components in the comprehensive decision-making process for a return to running.
Declarations

Ethics approval and consent to participate

All experimental protocols were approved by the local ethics committee of the University of Tokyo, and all participants gave their written informed consent.

Consent for publication

The participant depicted in the photos gave their written consent for publication.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgements

Not applicable.

Conflicts of interest
The authors declare no conflict of interest.

Funding statement: N/A


23. Stamenkovic A, Ting LH, Stapley PJ. Evidence for constancy in the modularity of trunk muscle activity preceding reaching: Implications for the role


FIGURE 1. Running and Nine Functional Tests. Participants independently performed running at a jogging pace and a walk. As a result, the average velocities for running and walk were noted to be $2.1 \pm 0.28$ m/s and $1.2 \pm 0.12$ m/s, respectively. In the single leg hops (SLH) tests, participants executed a forward jump at 30% (SLH30), 60% (SLH60), and 100% (SLH100) of their maximum distance. A single leg squat (SLS) was carried out with approximately 45 degrees of knee flexion. Hraise refers to the heel raise task, while stepup denotes a forward step up onto a 10 cm height box. Tbalance describes an exercise where the participant stands on the injured leg, forms a 'T' shape with the body, drives upward to a standing position, and then slowly returns to the 'T' position.

FIGURE 2. The percent of shared muscle synergies between running and each functional test. Median values are indicated as horizontal lines inside the boxes. The edges of the boxes represent the 25th and 75th percentiles. ($p = 0.000079$ for both groups, Friedman test one-way ANOVA). Statistically significant differences were observed between the following pairings: Walk and SLS ($p = 0.0038$, ES = 2.18), Walk and Hraise ($p = 0.0098$, ES = 2.05), SLH30 and SLS ($p = 0.0017$, ES = 1.56), SLH30 and Hraise ($p = 0.0039$, ES = 1.47), SLH and stepup ($p = 0.0072$, ES = 1.31), and Hraise and stepup ($p = 0.0176$, ES = 1.25).
FIGURE 3. Representative shared, running-specific, and test-specific muscle synergies identified by cluster analysis (muscle weighting components (W) and temporal pattern components (C)). The gray sections on the X-axis of C represent the temporal pattern components where a significant difference between running and functional tasks was observed (p < 0.05). RA: rectus abdominis, OE: oblique externus, ESL1 (erector spinae at L1), GM: gluteus maximus, Gmed: gluteus medius, BF: biceps femoris, ST: semitendinosus, TFL: tensor fasciae latae, ADD: adductor longus, RF: rectus femoris, VM: vastus medialis, VL: vastus lateralis, TA: tibialis anterior, MG: gastrocnemius medialis, SOL: Soleus (SOL) and POL: peroneus longus.
<table>
<thead>
<tr>
<th>run</th>
<th>walk</th>
<th>SLH30</th>
<th>SLH60</th>
<th>SLH100</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>W</td>
<td>[Graphs]</td>
<td>[Graphs]</td>
<td>[Graphs]</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>[Graphs]</td>
<td>[Graphs]</td>
<td>[Graphs]</td>
</tr>
</tbody>
</table>

**Shared**

**Run specific** or **Test specific** or **None**

<table>
<thead>
<tr>
<th>SLS</th>
<th>Ybalance</th>
<th>Hraise</th>
<th>stepup</th>
<th>Tbalance</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>C</td>
<td>[Graphs]</td>
<td>[Graphs]</td>
<td>[Graphs]</td>
</tr>
</tbody>
</table>

**Shared**

<table>
<thead>
<tr>
<th>Run specific</th>
<th>Test specific</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Graphs]</td>
<td>[Graphs]</td>
<td>[Graphs]</td>
</tr>
<tr>
<td></td>
<td>VAF values</td>
<td>Number of synergies</td>
</tr>
<tr>
<td>---------------------</td>
<td>------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Running</td>
<td>0.92 (± 0.01)</td>
<td>3.9 (± 0.31)</td>
</tr>
<tr>
<td>Functional tests</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walk</td>
<td>0.92 (± 0.01)</td>
<td>3.8 (± 0.63)</td>
</tr>
<tr>
<td>SLH30</td>
<td>0.92 (± 0.01)</td>
<td>4.3 (± 0.67)</td>
</tr>
<tr>
<td>SLH60</td>
<td>0.91 (± 0.01)</td>
<td>4.6 (± 0.51)</td>
</tr>
<tr>
<td>SLH90</td>
<td>0.91 (± 0.01)</td>
<td>5.1 (± 0.56)</td>
</tr>
<tr>
<td>SLS</td>
<td>0.92 (± 0.01)</td>
<td>2.4 (± 0.51)</td>
</tr>
<tr>
<td>Ybalance</td>
<td>0.92 (± 0.01)</td>
<td>3.6 (± 0.84)</td>
</tr>
<tr>
<td>Hraiese</td>
<td>0.92 (± 0.01)</td>
<td>2.3 (± 0.67)</td>
</tr>
<tr>
<td>Steuup</td>
<td>0.91 (± 0.01)</td>
<td>4.5 (± 0.97)</td>
</tr>
<tr>
<td>Tbalance</td>
<td>0.92 (± 0.01)</td>
<td>3.4 (± 0.69)</td>
</tr>
</tbody>
</table>