Please quote this SOP in your Methods.

The use of four limb hanging tests to monitor muscle strength and condition over time

SOP (ID) Number	DMD_M.2.1.005
Version	3.0
Issued	August 3rd 2011
Last reviewed	May 27th 2019

Author	George Carlson Department of Physiology, Midwestern University Glendale, Glendae, AZ, USA
Working group members	Maaike van Putten, (Leiden University Medical Center, Department of Human Genetics, Leiden, the Netherlands) Annemieke Aartsma-Rus, (Leiden University Medical Center, Department of Human Genetics, Leiden, the Netherlands)
SOP responsible	George Carlson
Official reviewer	Maaike van Putten

TABLE OF CONTENTS

1 OBJECTIVE

The four limb hang test uses a wire grid system to non-invasively measure the ability of mice to exhibit sustained limb tension to oppose their gravitational force. The procedure measures the 4 limb hang time in seconds as well as the minimal Holding impulse (Holding Impulse = Body mass x Hang Time) that is used to oppose the gravitational force. The test can be used to determine the natural course of neuromuscular disease or the efficacy of genetic or pharmacologic treatment approaches. In comparison to age-matched nondystrophic mice (wild type, C57/Bl10ScSnJ), dystrophic (*mdx*) mice exhibit a significantly reduced ability to oppose their gravitational force at ages of 1 to 4 months. Older *mdx* mice (6 to 14 months), however, do not exhibit a consistent reduction in Holding Impulse in comparison to nondystrophic mice (1). However, when the longest hanging time method is used (where mice hang up to 600 sec), differences between wildtype and *mdx* mice can be recorded up to the age of 18 months (2).

2 SCOPE AND APPLICABILITY

The four limb hang test is performed with mouse models of neuromuscular disorders to demonstrate neuromuscular impairment and motor coordination. It is an efficient and reliable outcome measure for the evaluation of effects of potential therapeutic compounds on muscle strength. However, due to the nature of the test, it is not possible to relate the outcome to a sole neuromuscular defect, or muscle.

The test is extremely easy to perform and inexpensive, although animals need to be supervised at all times, making the number of animals which can be simultaneously analyzed relatively small. Mice are clearly willing to perform the test since behavior indicates that they do not want to fall off the grid. However, some very young mice often try to avoid hanging by jumping from the grid, or climbing on the grid. Animal behavior, balance and weight can influence the test. The Holding Impulse is used as an attempt to correct for the negative effects of body mass on the Hang Time. However, it is always useful to correlate body mass and Holding Impulse to assess the potential influence of body mass on the Holding Impulse. If mice within a particular treatment group are equally able to oppose their gravitational force, then there should be no correlation between body mass and Holding Impulse (slope = 0; line A in Figure 1). If larger mice within a treatment group have a greater ability to sustain tension at levels equal to or larger than their gravitational force, then the correlation between body mass and Holding Impulse should be positive (Fig. 1, Line B).
If smaller mice have a greater ability, then the corresponding correlation should be negative (Fig. 1 Line C). Such comparisons are easy to accomplish, since the body masses and Hang Times are routinely recorded for each session. These correlations provide important information regarding the relative efficacy of a particular treatment in increasing body mass and in improving the ability to produce sustained tension in the limb flexors.

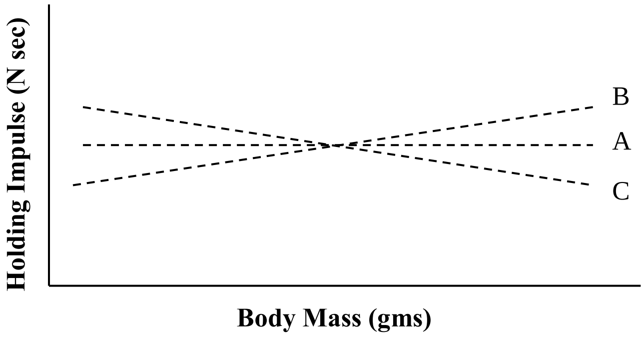

Figure 1. *Theoretical considerations regarding the relation between body mass and Holding Impulse. The line (A) indicates that mice with larger body masses have the same ability to produce sustained force to oppose gravity as mice with smaller body masses. (B) shows that larger mice have an improved ability to sustain limb tension to oppose gravity, and (C) shows that smaller mice have an improved ability.*

3 CAUTIONS

A high degree of standardization is needed to reduce inter-trial and inter-examiner variability. The sex, age and body weight of control and test animals should be similar. Also conditions of the room could influence the measurement and should be kept similar throughout the experiment (type of room, room temperature, room occupancy, time of the day, odors, etc). Inter-trial and inter-examiner reliability tests should be performed to assess and report the reliability of the test procedure used. This can be done by using two examiners to assess holding time and impulse on a sample of approximately 40 mice exhibiting a wide range of responses (1).

Soft bedding must be placed underneath the grid to break the fall and prevent the mice from harming themselves. We recommend a height of 35 centimeters. When it is set too high from the bedding, mice may harm themselves. However, the grid should be placed high enough to reduce the tendency of the mice to jump from the grid on purpose.

The investigators should be familiar with handling mice and be present throughout the experiment. Since the stopwatch needs to be stopped directly upon the beginning of each fall, only two mice can be measured simultaneously. Mice that jump off the grid should be immediately retested. If the mouse immediately jumps off the grid a second time, the results from that mouse should not be used. This behavior is observed in some young mice, but especially in strong wild type (C57BL/10ScSnJ) mice, during the first session. However, most mice comply with the test.

Advantages

The four limb hang test is a very inexpensive test since the equipment can be self constructed or one can use the top of a cage (tops of rat cages work well; Fig 2). Most animals are willing to perform the task, despite their physical condition.

Figure 2. *Mdx mouse hanging with four limbs on a grid.*

Disadvantages

No more than two mice can be tested simultaneously since constant supervision is necessary. Mice (often young wild types) that jump off the grid, or climb on top of the grid on purpose need to be returned to the grid by the investigator. The Holding Time and Impulse measures exhibit high variability, making it necessary to use large sample sizes and assess inter-examiner and inter-trial reliability (1).

4 MATERIALS

Grids can be easily homemade from aluminum mesh screens. One should pay attention to the size of the squares, which should not be too wide (1x1 cm square grid or smaller is suitable). The grid can either be a square (3, 4) or be put in a cup-like device in which the mesh screen (3.3 mm grid) is attached to the bottom of a round plexiglass cylinder (height 15 cm, internal diameter 12 cm, wall thickness 5 mm (1); Figure 3). The grid should not contain sharp things on which the mouse could harm himself and should be placed at least 35 cm. above a cage with sufficient bedding to ensure a soft landing. When placed too high, mice might be harmed upon landing, when placed too low, mice could let themselves fall on purpose and thereby avoiding hanging. A stopwatch is needed for accurate time recordings.

Mice as young as four weeks and as old as 18 months have been reliably evaluated with this test. Mice younger than four weeks may be too excitable to provide reliable values. No acclimatization to the test is needed, however, to limit stress to the animals, it is preferable that the mice have been handled on a regular basis before the test. It should also be kept in mind that differences in genetic background could influence the test. Therefore C57BL/10ScSnJ mice should be used as wild types for comparisons with *mdx* mice.

5 METHODS

Wire Screen Holding Test

The four limb hang test as described here was adapted from (5) and has been more thoroughly tested in a paper by Carlson et al (1). The method described here makes use of the cup-like device (Fig.3).

Figure 3. *Cup-like device used to assess Holding impulse. Left side shows cup upside down (bottom of cup at top) and right side shows the cup bottom tilted to expose the aluminum mesh screen. The mouse is placed upright on the mesh screen inside the cup before the cup is turned upside down.*

1. The mouse is placed inside at the bottom of the cup and is allowed to accommodate to this environment for 3 to 5 seconds before the cup is inverted and held at least 35 cm over a mouse cage containing 5 to 7 cm of soft bedding (e.g., wood chips). Each hang period must begin with all four paws of the mouse grasping the screen.

2. Typically, the mice lose their grasp of the mesh one or two paws at a time. Shortly before they fall, most mice usually have only 1 or 2 paws grasping the screen. The wire screen holding time (or "hang time" in seconds) is defined as the amount of time that it takes the mouse to fall from the inverted screen. The hang time is measured from the time the cup is inverted to the time that the mouse falls off the wire grid (determined visually and measured using a stopwatch). One could choose to either give the mice an unlimited amount of time to hang, or apply a fixed limit (e.g., a maximum of 600 seconds hanging).

3. The procedure may be repeated several times for each mouse with a rest interval between hang attempts of 2 – 3 minutes. The mouse body weight is obtained shortly before or after the test.

4. The Holding impulse associated with the holding test equals the hang time multiplied by the body weight (gm sec or Newtons sec; conversion factor - 9.806×10^{-3} Newtons/gm).

5. The determination of whether mice are given an unlimited amount of time to hang or are otherwise permitted to hang for a fixed time limit depends upon practical considerations and the behavior of the mice under investigation. If the mice under investigation often hang for prolonged periods (e.g., > 600 seconds), then a fixed limit should be imposed, and the

data should be analyzed in accordance with the protocol in (6). If the mice instead hang for shorter times, then the protocol identified in (7) should be employed.

6. When mice are given a fixed maximum time (e.g., 600 sec), mice that fall off the grid before the time limit will be directly given two more tries, while mice that hang for the 600 second limit will be placed back in the cage. In this case, either the maximum hang time observed or the mean of the observed hang times can be used as the dependent variable. When applying a fixed maximum time, the holding impulse cannot be used, since the unlimited maximum hanging time remains unknown. An exemplary video for this procedure has been generated (6).

7. When mice typically hang for shorter periods (< 600 sec), it is useful to give each mouse three hanging attempts without applying a set time limit, but with suitable rest periods (at least 2 minutes) between attempts. In this case, either the maximum hang time observed or the mean of the observed hang times can be used as the dependent variable. With no restriction to the hanging time, the Holding impulse should be assessed.

8. It is useful to perform inter-trial and inter-examiner reliability tests for the dependent variables identified in (6) and (7). These are accomplished by performing the procedure twice on a large sample of mice (e.g., 40 mice) using two investigators (Investigators A and B). During session 1, investigators A and B each examine two equal (e.g., 20 mice) and distinct subsets of the sample (e.g., Investigator A examines subset I and Investigator B examines subset II). During session 2, investigator A and B switch the subsets (Investigator A examines subset II, Investigator B examines subset I). A period of approximately 1 week is allowed between the two sessions. Inter-trial reliability is tested by correlating the results from each session for each mouse, regardless of investigator. Inter-examiner reliability is tested by correlating the results from each investigator for each mouse, regardless of session (1).

Inter-examiner (Fig 4A1-3) and inter-trial (Fig. 4B1-3) reliability determinations have been made for the protocol identified in (7) in which each mouse was given three trials per session. In this case, the dependent variables included the first Holding impulse (Fig. 4A1, A2), the maximum Holding impulse (Fig. 4B1, B2), and the mean of three Holding impulses (Fig 4C1,C2 (1). The correlation coefficients indicated in the insets to the figure represent the reliability for each assessment.

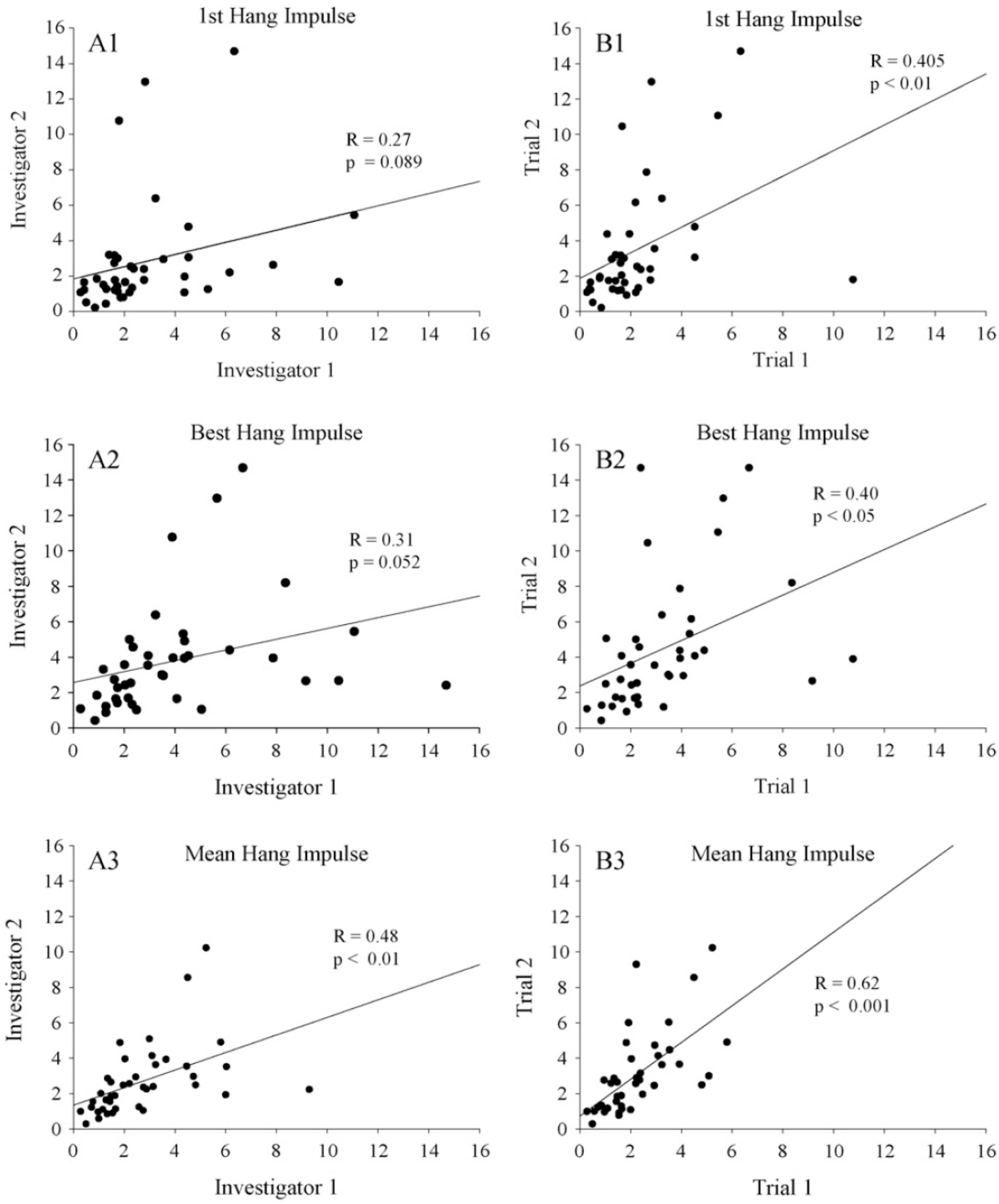

Figure 4. *Inter-examiner (A1 to A3) and inter-trial (B1 to B3) reliability determinations for the first (A1, B1), maximum (A2, B2), and mean of 3 Holding impulses in a sample of nondystrophic and mdx mice. Y axis units are in Newtons sec (from Carlson et al., 2010).*

Figure 5 shows the Holding impulses determined using the mean of 3 holding attempts on a large sample of nondystrophic and *mdx* mice. The results indicate a significant deficit in Holding impulse in *mdx* mice at ages less than 4 months. *Mdx* mice older than 6 months did not exhibit a consistent deficit in holding time (Carlson et al., 2010).

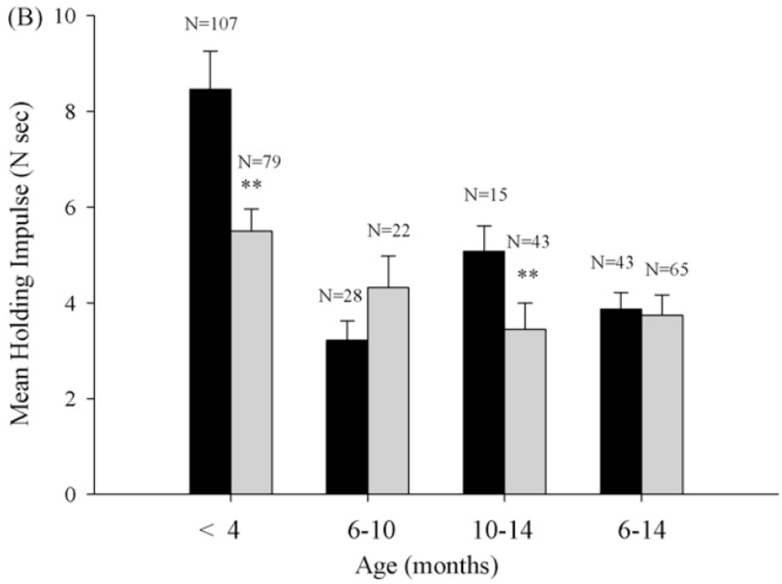

Figure 5. *Holding impulses determined over the first 14 months of the lifespan of nondystrophic (wild type, black histobars) and mdx (gray histobars) mice. Units are in Newtons sec. Note that the mdx mice at 6 to 10 months do not exhibit a deficit while those at 10 – 14 months do exhibit a deficit. Combining these results indicated that there was no consistent and significant deficit over the age of 4 months in the mdx sample. The results utilized the mean of 3 successive attempts using the protocol identified in (7) above (i.e., no fixed maximum holding time imposed). The double asterisks (**) indicate p<0.01 in comparison to the age-matched wild type mice (from Carlson et al., 2010).*

The four limb hang test has also been used to assess functional performance in several dystrophic mouse strains, with 0 (-/-), 1 (+/-) or 2 (+/+) utrophin alleles (Fig 6). Utrophin partly takes over the function of dystrophin in mice. Mice without utrophin (*mdx/utrn -/-*) are severely affected and die before the age of three months. These data have been obtained with a grid from a rat cage (Fig. 2) instead of the cup-like device and used the protocol indicated in (6) above (i.e., maximum hang time imposed). The mice had to hang for 600 seconds, when mice succeeded, they were placed back in their cage; otherwise, they were given two more tries. Data show that *mdx/utrn +/-* mice perform worse compared to *mdx/utrn +/+* mice. The performance of *mdx/utrn -/-* mice does not differ from that of *mdx/utrn +/+* and wild type (C57BL/10ScSnJ) mice, which might be because of their reduced body weight (4).

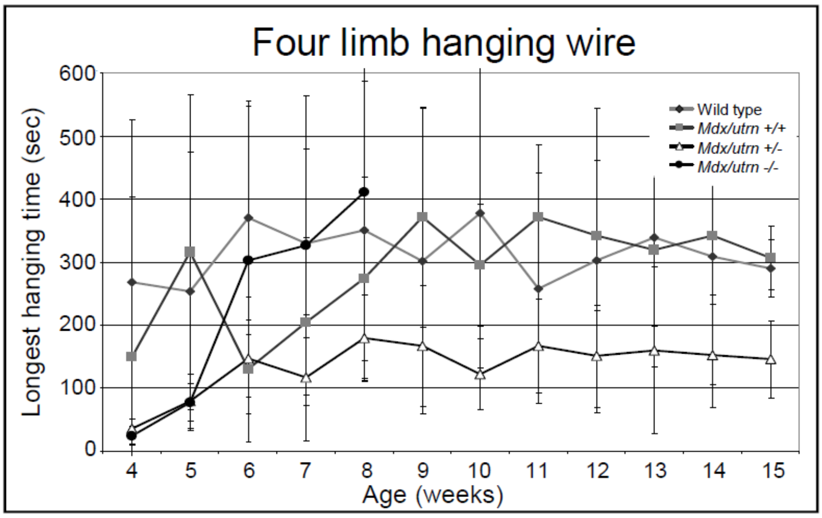

Figure 6. *Hanging time measures from different dystrophic mouse strains assessed weekly for a period of 12 weeks. Data resemble the longest hanging time out of three. Maximum allowed hanging time was 600sec. Mdx/utrn +/- mice perform worse than mdx/utrn +/+ and wild type mice (p<0.01). Mdx/utrn -/- mice died prematurely, but were able to perform the test (from van Putten et al, 2012).*

6 EVALUATION AND INTERPRETATION OF RESULTS

The four limb hanging test measures the ability of the mice to produce sustained tension in the limb musculature and should be distinguished from measures which determine phasic force generation in the limb musculature.

1. There is considerable variability in holding time. Many mice must be sampled in order to detect significant differences. In addition, differences in genetic background can influence the data.

2. The technique has been tested for reliability in populations of *mdx* and nondystrophic mice. Both inter-trial and inter-examiner reliability have been tested (1).

3. The hang time test is not a measure of phasic tension generation since it involves the maintenance of a minimum force required to oppose the mouse's gravitational force. The Holding impulse is a measure of that sustained force.

The hang test is easy to perform, applicable for both young and old mice and can be used in conjunction with other noninvasive measures of phasic tension generation to screen potential compounds for whole body efficacy, without influencing the natural course of the disease (3).

7 REFERENCES

1. Carlson CG, Rutter J, Bledsoe C, Singh R, Hoff H, Bruemmer K, et al. A simple protocol for assessing inter-trial and inter-examiner reliability for two noninvasive measures of limb muscle strength. J Neurosci Methods. 2010;186(2):226-30.

2. Kogelman B, Putker K, Hulsker M, Tanganyika-de Winter C, van der Weerd L, Aartsma-Rus A, et al. Voluntary exercise improves muscle function and does not exacerbate muscle and heart pathology in aged Duchenne muscular dystrophy mice. J Mol Cell Cardiol. 2018;125:29-38.

3. van Putten M, de Winter C, van Roon-Mom W, van Ommen GJ, t Hoen PA, Aartsma-Rus A. A 3 months mild functional test regime does not affect disease parameters in young mdx mice. Neuromuscul Disord. 2010;20(4):273-80.

4. van Putten M, Kumar D, Hulsker M, Hoogaars WM, Plomp JJ, van Opstal A, et al. Comparison of skeletal muscle pathology and motor function of dystrophin and utrophin deficient mouse strains. Neuromuscul Disord. 2012;22(5):406-17.

5. Hosaka Y, Yokota T, Miyagoe-Suzuki Y, Yuasa K, Imamura M, Matsuda R, et al. Alpha1-syntrophin-deficient skeletal muscle exhibits hypertrophy and aberrant formation of neuromuscular junctions during regeneration. J Cell Biol. 2002;158(6):1097-107.

6. Aartsma-Rus A, van Putten M. Assessing functional performance in the mdx mouse model. J Vis Exp. 2014(85).