



Published in final edited form as:

*Mech Ageing Dev.* 2019 June ; 180: 49–62. doi:10.1016/j.mad.2019.04.001.

## High intensity interval training improves physical performance in aged female mice: A comparison of mouse frailty assessment tools

**Kenneth Ladd Seldeen\***,

**Yonas Z. Redae,**

**Ramkumar Thiyagarajan,**

**Reem Nagi Berman,**

**Merced Marie Leiker,**

**Bruce Robert Troen**

Division of Geriatrics and Palliative Medicine, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo and Research Service, Veterans Affairs Western New York Healthcare System, Buffalo, NY, USA

### Abstract

Frailty syndrome increases the risk for disability and mortality, and is a major health concern amidst the geriatric shift in the population. High intensity interval training (HIIT), which couples bursts of vigorous activity interspersed with active recovery intervals, shows promise for the treatment of frailty. Here we compare and contrast five Fried physical phenotype and one deficit accumulation based mouse frailty assessment tools for identifying the impacts of HIIT on frailty and predicting functional capacity, underlying pathology, and survival in aged female mice. Our data reveal a 10-minute HIIT regimen administered 3-days-a-week for 8-weeks increased treadmill endurance, gait speed and maintained grip strength. One frailty tool identified a benefit of HIIT for frailty, but many were trending suggesting HIIT was beneficial for physical performance in these mice, but the 8-week timeframe may have been insufficient to induce frailty benefits. Finally, most frailty tools distinguished between surviving or non-surviving mice, whereas half correlated with functional capacity measured by nest building ability, and none correlated with underlying pathology. In summary, this study supports the ongoing development of mouse assessment tools as useful instruments for frailty research.

### Keywords

Exercise; Physical performance; Functional capacity; Aging; Survival

---

\*Corresponding author at: 875 Ellicott Street, CTRC Room 8030A, Buffalo, NY 14203, USA. seldeen@buffalo.edu (K.L. Seldeen). Author contributions

KLS, YZR, RT, and RNB performed experiments. KLS and BRT contributed to experimental design and analyzed data as well as prepared and wrote manuscript.

Declaration of interest

The authors of this manuscript declare no conflicts of interest.

## 1. Frailty introduction

Frailty is a syndrome characterized by poor resilience that increases susceptibility to disability, loss of independence, hospitalization, and mortality (Fried et al., 2001; Rockwood et al., 2005). Frailty is prevalent and increases with age, rising from 10% in individuals 65 and over to 50% in those over the age of 85 (Collard et al., 2012; Gonzalez-Vaca et al., 2014). Clinical efforts to understand and treat frailty were aided by the introduction of frailty quantification tools like the Fried physical frailty phenotype (Fried et al., 2001) and the Rockwood et al. deficit accumulation frailty index (Rockwood et al., 2005). The former establishes frailty as exhibiting 3 or more traits including unexpected weight loss and low grip strength, activity levels, endurance, or gait speed (Fried et al., 2001), while the latter generates a continuous score from a wide range of parameters stemming from symptoms, signs, illnesses, and disabilities (Rockwood et al., 2005). Additionally, both frailty assessment tools are predictive for health outcomes that include functional capacity, disability, and mortality (Li et al., 2015).

### 1.1. Animal frailty assessment tools

Animal models provide the framework to reduce confounding life-style and genetic factors, examine the impacts of genetic modification, and observe the impacts of aging in a timeframe that is amenable to laboratory science. These advantages have ignited efforts by several research groups to establish animal frailty assessment tools that likewise are capable of quantifying the condition and predicting health outcomes and identify therapeutic benefits. A total of seven mouse frailty tools have been established to date.

**1.1.1. Fried physical phenotype based mouse assessment tools**—Liu and Graber et al. established the first physical phenotype based animal frailty tool with 4 parameters that included an inverted grid hang for grip strength, activity wheel usage for activity levels, rotarod maximum speed for gait speed, and a combination of grid hang and rotarod for endurance (Liu et al., 2014). Kane et al. subsequently appended mouse weight to this method to complete the 5 physical phenotype parameters similar to the Fried assessments in humans (Kane et al., 2017a). Next, Gomez-Cabrera et al., established the Valencia score for frailty assessment in mice based on weight loss, treadmill performance for endurance and gait speed, grip strength, and balance beam coordination for activity levels (Gomez-Cabrera et al., 2017). Both Liu and Graber and the Valencia score tools were predictive of age-associated decline (Liu et al., 2014; Gomez-Cabrera et al., 2017), while the Valencia score was further shown to predict lifespan (Martinez de Toda et al., 2018). Finally, our research group (Seldeen et al., 2018a) established a frailty assessment tool for mice using weight loss, treadmill endurance, gait speed, grip strength, and open field activity (Seldeen et al., 2018a).

**1.1.2. Rockwood frailty index based mouse assessment tools**—Parks and Howlett et al., led the first effort to create a deficit accumulation based frailty tool for mice (Parks et al., 2012). The tool included a total of 31 parameters drawing from an assessment of open field activity, serum metabolites, hemodynamic parameters, and body composition determined using dual X-ray absorptiometry (Parks et al., 2012). Although the

frailty tool was predictive of age associated changes, Whitehead and Howlett et al. made further efforts to create a protocol that was more universal and applicable across multiple laboratories (Whitehead et al., 2014). This effort led to the mouse clinical frailty index (C-FI), a 31-parameter index derived from a physical inspection of the animals, and requiring only a noise generating device and infrared thermometer of negligible costs (Whitehead et al., 2014). The C-FI has since been shown to identify the impacts of age (Whitehead et al., 2014), predict lifespan (Kane et al., 2017b), and assess the impacts of therapeutic interventions (Kane et al., 2016; Keller et al., 2018).

**1.1.3. Other frailty assessment tools**—Two other frailty assessment tools have been published. The first by Antoch et al., developed the physiological frailty index (PFI) for mice, which is comprised of measures of body weight, grip strength, blood cell composition (white and red blood cell counts and differentials), plasma concentrations of CXCL1, triglycerides, and glucose, as well as analysis of the cardiovascular system (systolic, diastolic, blood pressure, heart rate, tail blood flow, and tail blood volume) (Antoch et al., 2017). The authors show this frailty tool identifies higher rates of frailty in older mice and during obesity, as well as the benefits of therapeutic interventions (Antoch et al., 2017). The second, the frailty intervention assessment value (FIAV), may also be a useful tool for the analysis of frailty in mice, although this tool was first introduced by Graber et al. as a human equivalent Barthel index (Graber et al., 2015).

## 1.2. Exercise for frailty

The evidence for the benefits of exercise for frailty in humans is strong as four studies (Cameron et al., 2013; Manas et al., 2018; Rogers et al., 2017; Yamada et al., 2012) and 1 review (Apostolo et al., 2018) found that moderate exercise regimens reduced frailty. Despite the benefits of exercise, only 25% of individuals meet the recommended 150 min/week of moderate intensity exercise with strength training (Clark et al., 2017), and participation declines with age as only 8% of those over 75 meet these activity recommendations (Clark et al., 2017). Lack of time and length of exercise sessions are cited as the most common barriers to participation (Justine et al., 2013; Schutzer and Graves, 2004), thus high intensity interval training (HIIT), with short sessions and less time commitment, is emerging as a strategy to delay the onset of frailty. Fiatarone et al. (1990) pioneered the use of HIIT as an intervention for the very old in a study involving nonagenarians (age  $90 \pm 1$  years) and found excellent tolerance while improving strength and gait speed over 10 weeks. HIIT has also been shown to be safe and effective while improving fitness, insulin sensitivity, muscle hypertrophy and power, cardiovascular function, and quality of life in individuals > 65 years of age (Bell et al., 2015; Bruseghini et al., 2015; Knowles et al., 2015; Sculthorpe et al., 2015). Finally, it was recently demonstrated that HIIT, as a component of a larger therapeutic strategy, reduced frailty in individuals over the age of 75 (Losa-Reyna et al., 2019).

## 1.3. Animal frailty and exercise and study purpose

Several attempts have been made to apply frailty assessment tools to observe whether exercise is beneficial for aged mice. The first application by Graber et al. used the frailty tool from Liu and Graber et al. in addition to the FIAV tool to identify the benefits of

voluntary wheel running (Graber et al., 2015). Gomez-Cabrera et al. further demonstrated the benefits of voluntary wheel running for frailty in mice using the Valencia score (Gomez-Cabrera et al., 2017). Finally, our research group applied our frailty tool and a modified version of FIAV to demonstrate four months of a thrice weekly, 10-minute HIIT regimen reduces frailty in 24-month old aged C57BL/6J male mice (Seldeen et al., 2018a). However, to date there have been no studies that examine how these various frailty tools compare in examining the benefits of exercise, particularly the inclusion of non-frailty phenotype tools like the C-FI. Therefore, here we compare the Seldeen, Liu & Graber, Valencia Score, C-FI, and FIAV frailty assessment tools in the context of our HIIT protocol in aged mice. Additionally, this was performed using females as the impacts of HIIT exercise have not been previously examined in female mice, and female humans are underrepresented in HIIT clinical studies. Our data reveal HIIT exercise induces multiple physical performance benefits for female mice, although only one frailty tool identified a reduction in frailty. However, the frailty tools are able to predict functional capacity and mortality. These findings support the use of these tools for the ongoing study of the mechanistic underpinnings of frailty in animal models and lay a foundation for better understanding of human frailty, as well as potentially beneficial interventions.

## 2. Materials and methods

### 2.1. Animals

C57BL/6 J female mice were bred and aged in house. The mice contained a  $SIRT1^{loxP-exon4-loxP}$  mutation (Cheng et al., 2003). However we did not induce the conditional mutation, no differences were observed in Sirt1 protein size or levels relative to wild type mice, and these mice were phenotypically identical to wild type mice. These mice were used because of the availability of large numbers of aged females in our colony. Mice were bred and aged in the Optimice ventilated rack system (Animal Care Systems, Inc., Centennial, CO, USA) until 21 months of age, at which point 31 mice were transferred to traditional wire top shoebox cages containing 2–3 mice per cage. Food and water were provided *ad libitum* and lighting was on a standard 12 h on/off cycle. Mice were provided facility chow (Teklad Global 18% #2018XS by Envigo, Madison Wisconsin USA). Mice also received veterinary support as needed including analgesics, fluid injection, and treatment for malocclusions. Veterinary staff were blind to the group designations of the mice. Acclimation to physical performance and frailty assessments was initiated at 23 months of age, followed by baseline assessments at 24 months of age. The exercise phase lasted a total of 8 weeks, with final assessments performed the week following. Four animals died prior to the exercise phase of the experiment, therefore the remaining mice ( $N = 28$ ) were randomly assorted by body weight into either sedentary (SED) or HIIT groups. However, three mice died during the exercise phase of the experiment (all from the sedentary group), therefore final numbers were SED ( $n = 11$ ), HIIT ( $n = 14$ ), and deceased ( $n = 7$ ). Body weight was measured weekly. The week following endpoint assessments, mice were culled and a post-mortem inspection was conducted for the presence of tumors or other tissue abnormalities. To quantify these observations, health scores were given as 1 for tumors, enlarged organs or other serious abnormalities, 2 for minor issues, and 3 for otherwise non-remarkable. All studies and experimental protocols were approved by and in

compliance with guidelines of the University at Buffalo and VA Western New York Animal Care and Use Committees.

**2.1.1. Physical performance assessments**—Baseline and endpoint assessments were performed at the same time of day and in the same order during the week.

Additionally, the assessments were performed by the same investigator who was blind to the group designations of the animals.

**2.1.2. Flat continuous and uphill interval treadmill assessment**—The protocol for treadmill assessments was described previously (Seldeen et al., 2018a) and in detail (Seldeen et al., 2018b). Briefly, animals were trained and acclimated to a mouse treadmill with shock detection (Columbus instruments, Columbus, OH, USA) at one month and two weeks prior to baseline assessments. For flat continuous treadmill assessment (treadmill belt at 0°), the speed of the belt starts at 5 m/min and accelerates 1 m/min every two minutes until exhaustion as defined as 10 visits to a shock grid or 20 total shocks. For uphill interval treadmill assessment (belt at 25°), the belt starts at 5 m/min for 30 s, and then provides incrementing test stages of 20 s (at 6, 8, 10 m/min, then +1 m/min thereafter), with 20 s active recovery stages between. Exhaustion is defined as 5 visits to the shock grid or 10 total shocks. Due to the diversity in body sizes, shock intensity was individually assigned to mice based upon body weight at the time of the experiment. Mice less than 40 g received shocks at an intensity of 0.72 mA, from 40 to 50 g 0.84 mA, and greater than 50 g, 0.97 mA.

**2.1.3. Grip strength**—Grip strength was assessed as previously described (Seldeen et al., 2018a). Briefly, the best three of five trials are averaged with each consisting of three attempts whereby the mouse is held by the tail and placed with all four limbs on a grid attached to a force meter (Columbus instruments). The mouse is pulled until loss of grip, with ~1 s of rest between attempts and 10 s between trials. Mice were given 5 trials 1-month prior to baseline assessments to acclimate mice to the device.

**2.1.4. Activity monitor**—Open field activity was measured as described previously (Seldeen et al., 2018a) using an 8-station activity monitor (Med Associates Inc., Fairfax, VT, USA). Briefly, mice were placed into a clear plastic arena lined with infrared beams that identify the position of the mouse. After 30 min, total distance travelled, rearings, quadrant crossings, and average speed were calculated by the equipment software.

**2.1.5. Rotarod treadmill**—Rotarod assessment was performed as described previously (Seldeen et al., 2018a). Briefly, one month prior to baseline assessments, mice were acclimated to the rotarod device (Med Associates Inc.) by providing three trials whereby the speed increases from 2 to 20 RPM over 5 min. For assessment the rotarod latency was measured as the average of 3 trials as the device accelerates from 4 to 40 RPM over 5 min.

**2.1.6. Gait speed**—Gait speed was assessed as described previously (Seldeen et al., 2018a; Justice et al., 2014). Briefly, one month prior to baseline assessments, mice were acclimated to a dark “safe house” at the end of the gait speed device (Custom CoLaborators, LLC., Clarence, NY, USA). Next, mice were assessed as the best 2 of 3 trials whereby a

mouse is placed at the start of the device and timed until the mouse reaches the front of the safe house.

**2.1.7. Activity wheel monitoring**—One month prior to baseline, at baseline, and at endpoint, mice were provided activity wheels (Med Associates Inc.) in the home cages starting at 5 P M. local time and remaining for 48 h. The activity wheels transmit total revolutions wirelessly to equipment software on a central computer that calculates total revolutions and distance travelled (in kilometers).

**2.1.8. Grid hang (inverted cling)**—Mice are initially acclimated to a square wire grid (30.5 cm × 30.5 cm) as it is inverted over the home cages. Two weeks prior to baseline to further acclimate, at baseline, and at endpoint, mice are given the best of two trials whereby the mouse is inverted atop a 45.5 cm tall box and timed until fall. For each trial the mice are given three attempts to reach a minimum of 15 s, and no less than 10 min rest between trials. Trials were stopped if the mice remained on the grid for 300 s.

## 2.2. Frailty indices

We evaluated a total of 6 frailty assessment tools, with four based upon the model of the Fried frailty phenotype scale (i.e., five parameters with having 3 or more classified as frail, and 1 or 2 being pre-frail), one based upon deficit accumulation (i.e., one point given per presence of frailty criteria with a final score being the sum of all criteria divided by the total number of criteria), and one based upon the Barthel index (Graber et al., 2015). We did not include the assessment tools described by Parks et al. (2012) and Antoch et al. (2017) due to not having access to the specialized equipment and methodologies used in these approaches.

**2.2.1. Seldeen/Seldeen50**—The Seldeen frailty score was determined as described in Seldeen et al. (2018a), whereby a mean and standard deviation were calculated for the SED mice for grip strength, flat continuous treadmill, activity monitor crossings, and gait speed at baseline and again at endpoint. Cutoff for each parameter was determined as being below 1.5 standard deviations of the mean. Cutoff for weight was also determined as the loss of weight greater than 5% in one week prior to the start of the exercise phase and prior to endpoint assessments. The Seldeen50 frailty score was created to minimize the animal variability that resulted in one parameter having a cutoff score below zero and therefore not being usable. To adjust for this the mean and standard deviation were generated using only the 50% of the SED mice (total 6) closest to the group mean.

**2.2.2. Liu&Graber50**—The Liu&Graber50 frailty score was calculated based upon the strategy described in Liu et al. (2014), however, animal variability caused multiple parameters to have cutoff values below zero. Therefore, mean and standard deviations were generated from the baseline assessment of the 50% of surviving mice closest to the group mean, and used to establish cutoffs 1.5 SD below the mean of each activity for both baseline and endpoint. Parameters include grip grid latency, speed in RPM of rotarod at the time of fall, activity wheel usage, and the addition of grip grid latency with rotarod latency. Additionally, a weight parameter cutoff of 0.8 SD from the group mean was included as described (Kane et al., 2017a).

**2.2.3. Valencia score**—The Valencia score was calculated based upon methodology in (Gomez-Cabrera et al., 2017). Briefly, the body weight parameter was determined as a loss of weight greater than 5% in the one month prior to the start of the exercise phase and prior to the start of endpoint assessments. A cutoff value was set as the lowest 20% of mice for each parameter including flat continuous treadmill time on belt, the speed of the treadmill at the point of exhaustion, and grip strength. Additionally, the original protocol uses a wire hang test as a parameter for activity, however we have substituted the grid hang assessment with a similar cutoff time of 60 s.

**2.2.4. Clinical Frailty Index (C-FI)**—The 31-parameter frailty index score was determined as described in Whitehead et al. (2014), including the use of a Infrascan surface infrared thermometer (La Crosse Technology, La Crosse, WI, USA) for body temperature calculation and a dog trainer noise clicker ([Amazon.com](https://www.amazon.com)) for hearing determination. The tumor category was removed from the analysis as mice identified with tumors were not found to have tumors in the post-mortem analysis. Final score was therefore the sum of the scores of all parameters divided by 30. Additionally, as a deviation from the methodology used in Whitehead et al. (2014), we determined body weight and body temperature scores using mean and standard deviation data from the entire cohort at baseline and at endpoint. The final score for each mouse was derived from the average of three independent investigators.

**2.2.5. Frailty intervention assessment value (FIAV)**—The FIAV calculation was performed as described previously (Seldeen et al., 2018a; Graber et al., 2015). Briefly, the parameters for FIAV include weight loss 1-week prior, flat continuous treadmill time on belt, grip strength, gait speed, and activity monitor crossings. To generate a Z-score for each mouse, the value for each parameter is subtracted from the baseline mean of all mice then divided by the standard deviation generated from all mice. The scores were then added together to generate FIAV1. At endpoint the process was repeated, using endpoint values for each mouse with the baseline mean and SD to generate FIAV2. Total FIAV was calculated by subtracting FIAV1 from FIAV2.

### 2.3. Nest building evaluation

To evaluate nest building, a total of 15 g of paper nesting material and a cotton nesting material pad (termed a nestlet) were placed in fresh cages to allow individual housing. After two nights, the overall quality of the nest was rated from 0 (no nest) to 5 (nest with four walls and covering) according the protocol outlined by Gaskill et al. (2013). Additionally, the amount of material collected and used in the nest as well as the use of the cotton nestlet was estimated as a percent of total. The final score for each mouse was derived from the average of two independent investigators.

### 2.4. High intensity interval training (HIIT)

We administered an individually tailored HIIT regimen as described previously (Seldeen et al., 2018a), and as described in detail (Seldeen et al., 2018b). The HIIT protocol was given 3-days-a-week and consisted of a 10-minute program (3minutes at a “base” speed, followed by 3 intervals of 1-minute “sprint” speed/1-minute “base” speed, and a final segment

that accelerates from “sprint” to “dash” speed over 1-minute). Mice were designated into intensity groups with defined “base”, “sprint”, and “dash” speeds based upon baseline uphill interval treadmill performance (Seldeen et al., 2018a, b). The three speeds were increased by 1 m/min every two weeks and individual mice were also switched between intensity groups depending on performance. Sedentary animals were taken in cages and placed beside the treadmill for one training session at each exercise time point.

## 2.5. Statistics

Statistical analysis was performed using XLStat statistical software (Addinsoft, New York, NY). A Student’s unpaired *t*-test was used for all comparisons between surviving SED and HIIT groups as well as for comparisons between surviving SED and deceased mice. A Student’s paired *t*-test was used for within group comparisons between baseline and endpoint. Pearson correlation analysis was used to examine relationships between frailty assessment tools, between frailty tools and nest building outcomes, and between frailty tools and days of survival. The cut-off for significant comparisons was  $p < 0.05$ . All data are presented as mean  $\pm$  standard deviation.

## 3. Results

### 3.1. High intensity interval training improves physical performance in aged female mice

We previously identified a 10-minute individually tailored uphill treadmill HIIT regimen that improves physical performance and reduces frailty in aged male mice (Seldeen et al., 2018a). To determine if the benefits of this exercise strategy extended to female mice, we assessed physical performance in sedentary (SED) and HIIT exercise mice before and after an eight-week regimen. Interestingly, we did not observe a difference in body weight due to HIIT exercise, although both groups lost substantial weight during the experimental time course (Fig. 1A, B). Consistent with this observation, quantitative magnetic resonance analysis of body composition also revealed no differences in percent body fat and absolute lean mass in HIIT mice as compared to SED mice (Fig. 2A and B). However, analyses using Student’s paired *t*-test to examine changes within group identified decline in body fat % within the HIIT group ( $p = 0.0299$ ), but not the SED group ( $p = 0.0552$ ), and neither showed decline in absolute lean body mass (SED:  $p = 0.59$  and HIIT:  $p = 0.26$ ) during the 8 weeks. Interestingly, bone mineral density (BMD), as measured by dual X-ray absorptiometry, was greater in HIIT mice at endpoint (SED  $49.4 \pm 4.6$  mg/cm<sup>2</sup> versus HIIT  $53.5 \pm 3.8$  mg/cm<sup>2</sup>,  $p = 0.0467$ ,  $n = 8$  and  $11$ , respectively, Fig. 2C), however BMD was not examined at baseline.

We also found that uphill interval treadmill performance, a parameter we used to determine individualized exercise intensities for these mice, increased from baseline in HIIT mice, but not SED mice (maximum speed: SED baseline  $16.9 \pm 3.8$  m/min to  $16.5 \pm 5.0$ ,  $p = 0.77$  versus HIIT  $14.6 \pm 4.3$  m/min to  $17.6 \pm 5.5$  m/min,  $p = 0.0025$ , Fig. 1C). Furthermore, HIIT mice exhibited increased flat continuous treadmill performance and gait speed, as well as did not show the decline in grip strength observed in SED mice (Table 1). Both groups exhibited greater open field activity and we did not observe statistically significant changes in either group with regard to rotarod or grid hang performance.



### 3.2. FIAV but not other frailty assessment tools detect benefits of HIIT exercise for frailty; only C-FI detects increasing frailty with aging

We further assessed how the 8-week HIIT regimen affected frailty status in the mice using several recently published frailty assessment tools (Fig. 3). As we previously found that HIIT reduced frailty in aged male mice, we initially applied the protocol described in Seldeen et al. (2018a), however, we did not observe an impact on frailty status associated with HIIT (Fig. 3A). However, the present cohort exhibited greater variability in performance than our previously published cohort, resulting in one parameter with a below zero cut-off score (activity as defined by open field crossings), preventing the use of this parameter as a frailty criterion. To adjust for this variability, a new tool was created whereby for all parameters the mean and cutoff values were established using the 50% of mice closest to the mean (Seldeen50). Using this rule we identified 3 mice as frail in the HIIT group at baseline and 2 at endpoint, yet still did not observe statistically significant improvement due to HIIT (Fig. 3B).

Next, we examined frailty tools published by Liu & Graber, et al. (Liu et al., 2014) and the Gomez-Cabrera et al. Valencia score (Gomez-Cabrera et al., 2017). Similar to the difficulties experienced with the Seldeen tool, we found many of the parameters in Liu & Graber exhibited cut-offs below 0. We therefore applied the new rule of using the 50% of mice closest to the group mean to establish new cut-offs, and implemented a criteria for body weight as described (Kane et al., 2017a) to create Liu&Graber50 (Fig. 3C). Using Liu&Graber50 we identified 5 of the 14 mice as frail at baseline and none at endpoint. However, we did not find a statistically significant in the change of frailty scores in the HIIT group (paired *t*-test  $p = 0.09$ ). The Valencia score (Fig. 3D) was not subject to the issues regarding variability, however we substituted the grid hang test for the wire hang test, used for the activity parameter, due to lack of equipment. The Valencia score assessment tool identified 6 mice as frail in the HIIT group at baseline and 4 at endpoint, but likewise we did not observe a statistically significant benefit of HIIT exercise ( $p = 0.09$ ). In comparing the change in frailty for each of the tools (Fig. 4), we note similarities across the tools suggesting HIIT is beneficial for frailty, yet only the FIAV tool detected a statistically significant change from baseline (SED:  $0.1 \pm 3.4$  versus HIIT:  $3.1 \pm 2.3$ ,  $p = 0.0145$ , Fig. 4F). In addition, we also identified correlation between all of the Fried physical frailty based assessment tools at baseline, and Liu&Graber50 and the Valencia score at endpoint (Table 2).

Our analysis also included the mouse clinical frailty index (C-FI) (Whitehead et al., 2014). The method includes 31 parameters obtained through clinical inspection of the mice. However, a category for the presence of tumors was excluded as tumors identified or not identified during the post-mortem inspection of the mice were not consistent with endpoint C-FI scores for that parameter. Although our data showed no effect of HIIT on C-FI scores, the C-FI was the only of the tested frailty assessment tools to observe age-associated decline in frailty scores (SED baseline:  $0.08 \pm 0.04$  versus endpoint  $0.12 \pm 0.07$ ,  $p = 0.046$ , Fig. 3E). Like the C-FI, the frailty intervention assessment value (FIAV) converts multiple parameters into a continuous variable. Using FIAV, which unlike the other tools where higher scores denote less frailty, we identified a statistically significant improvement in frailty scores in

the HIIT group (HIIT baseline:  $0.3 \pm 2.4$  versus endpoint  $3.4 \pm 1.7$ ,  $p = 0.0002$ , Fig. 3F). FIAV was also the only frailty tool to detect a difference in the change between the two groups (SED:  $0.1 \pm 3.4$  versus  $3.1 \pm 2.3$ ,  $p = 0.0145$ , Fig. 4F). Although FIAV was well correlated to the other frailty tools at baseline, the C-FI was correlated only to the Seldeen frailty tool at baseline (Table 2).

### 3.3. Most, but not all, frailty assessment tools predict premature mortality in mice

Following the completion of baseline frailty assessments we identified 4 mice that died prior to the exercise phase (and prior to assignment to HIIT or SED groups) of our experiment and another 3 mice that died before endpoint assessments (all from the SED group). We then compared baseline frailty status of the deceased mice ( $n = 7$ ) to the surviving sedentary mice ( $n = 11$ ) and HIIT mice ( $n = 14$ ) (Fig. 5), and found statistically significant differences between SED and deceased mice for 4 frailty tools including Seldeen ( $p = 0.0078$ , Fig. 5A), Seldeen50 ( $p = 0.0365$ , Fig. 5B), Valencia score ( $p = 0.0366$ , Fig. 5D), and FIAV ( $p = 0.0056$ , Fig. 5F). Graber&Liu50 was found to be trending ( $0.078$ , Fig. 5C), however, this frailty tool had an  $n$  of 6 as one mouse died before the rotarod test and thus was not included in the group. The C-FI frailty score was also trending higher in the deceased group ( $p = 0.12$ , Fig. 5E). Within the deceased mice, none of the frailty tools exhibited correlation to the number of days survived, while FIAV was closest to significance ( $r = 0.62$ ,  $p = 0.13$ ). Of all physical performance parameters, only gait speed was significantly lower in deceased mice (SED  $31.0 \pm 5.9$  cm/s versus deceased  $22.7 \pm 6.7$  cm/s,  $p = 0.0260$ , Table 3), while body temperature was the only parameter correlated to the number of days survived in the deceased mice ( $r = -0.84$ ,  $p = 0.0174$ , Table 3).

### 3.4. HIIT enhances functional capacity as determined by nest building ability in mice; Seldeen, Valencia score, and FIAV predict functional capacity

The activities of daily living (ADLs) and instrumental activity of daily living (IADLs) are important parameters in geriatric medicine to assess independence and functional capacity (Gulley et al., 2018). Likewise, nest building has been likened to a mouse IADL (Gaskill et al., 2013; Deacon, 2012). Deacon, 2012 We therefore compared nest building in mice before and after the 8 week HIIT exercise regimen (Fig. 6). Our data show HIIT mice produced higher quality nests than SED mice (nest score: HIIT  $4.0 \pm 0.8$  versus SED  $3.0 \pm 1.4$ ,  $p = 0.0304$ , Fig. 6C). Consistent with the improved nests, HIIT mice also used more paper nesting material (HIIT  $82.9 \pm 15.6\%$  versus SED:  $45.4 \pm 43.6\%$ ,  $p = 0.0064$ , Fig. 6D), and did not display the decline in the use of the cotton pad nestlet observed in the SED mice (paired  $t$ -test HIIT  $p = 0.2878$ , SED  $p = 0.0330$ , Fig. 6E). We further assessed whether the frailty assessment tools correlated with nest building scores across both surviving and deceased mice (Fig. 6F–H). At baseline, we found the Seldeen, Valencia score, and FIAV frailty tools had a statistically significant relationship with nest building scores ( $p = 0.0127$ ,  $0.0264$ , and  $0.0256$ , respectively). Interestingly, none of the frailty tools correlated with nest building at endpoint, with FIAV showing the closest correlation ( $p = 0.10$ ). Nest building did not predict survival in these mice (SED:  $3.4 \pm 1.5$  versus deceased  $2.4 \pm 1.9$ ,  $p = 0.25$ ,  $n = 7$ ).

### 3.5. Frailty assessment tools fail to correlate with post-mortem analysis

Following endpoint assessments, we performed a post-mortem inspection of major organs and tissues to better understand the relationship between frailty scores and underlying pathology. We quantified our observations by assigning a score of “1” for the presence of tumors, enlarged organs, or other serious abnormalities, a “2” for minor issues, and a “3” for otherwise unremarkable/healthy. Our data revealed that there were no differences in the underlying pathology between HIIT and SED mice ( $p = 0.3585$ , Fig. 7A). Next we correlated post-mortem scores with endpoint frailty scores and surprisingly did not find significant correlation with any frailty tool. Fig. 7B shows the scale that exhibited the greatest  $R^2$  value (Seldeen50) of only 0.01, with a p-value of 0.67.

## 4. Discussion

Here we demonstrate a 10-minute, 3-days-a-week, HIIT regimen increased treadmill endurance and gait speed, as well as mitigated grip strength decline in 24-month old aged female mice over an 8week period. These findings are consistent with our previous study wherein HIIT significantly enhance physical performance in older male mice and also reduced frailty (Seldeen et al., 2018a). Interestingly we find that none of the Fried based frailty tools identified benefits of exercise, although they were trending towards significance (Seldeen,  $p = 0.16$ , Liu&Graber50,  $p = 0.09$ , and Valencia,  $p = 0.09$ ). The 8-week timeframe may have been insufficient for those tools to detect an impact. Two mouse studies in addition to ours have utilized frailty tools and found exercise to be beneficial (Gomez-Cabrera et al., 2017; Graber et al., 2015). Several human clinical studies have also demonstrated that exercise mitigates frailty (Cameron et al., 2013; Manas et al., 2018; Rogers et al., 2017; Yamada et al., 2012).

However the FIAV tool indicated that HIIT had a significant and beneficial impact in reducing frailty ( $p = 0.0002$ ), and this was also consistent with our previous study in male mice (Seldeen et al., 2018a). One aspect that may have resulted in this outcome is FIAV uses a continuous scoring system that may be more responsive to improvement than the defined cutoffs used in the Seldeen, Liu&Graber, and Valencia Score tools. As an example, three HIIT mice improved greater than 5-fold in treadmill performance, yet remained slightly below the endpoint cutoff for the endurance parameter of the Seldeen50 frailty tool. Additionally, the sensitivity of the Liu&Graber and Seldeen frailty tools may have been diminished due to the large variability in animal performance. In fact, the Liu&Graber tool was not usable as originally described (Liu et al., 2014), and both tools appeared to have increased sensitivity by establishing cut-off scores based upon using only the 50% of mice closest to the group mean. We note that the Valencia score was not as susceptible to mouse variability as the other Fried based frailty tools due to automatic assignment of the cut-off as the lowest 20% of mice. The correlation between the Fried frailty tools at baseline and endpoint suggests that either system is sufficient for defining cutoffs. One exception is the original Seldeen frailty tool had poor correlation with the other tools at endpoint, which we believe is due to very few animals being identified as pre-frail or frail at endpoint using this tool.

The C-FI was not able to discern a reduction in frailty score in response to HIIT. Perhaps a reason was that only a handful of the 31 frailty (or 30 in our study) criteria may be sensitive to the benefits of exercise (e.g., grip strength, gait speed, body condition) or perhaps the mice did not exhibit sufficient frailty to observe an impact. Further studies using older mice, a lengthier HIIT protocol, or other exercise modalities would aid in supporting this point, particularly as the C-FI frailty score did increase in the sedentary group, but did not increase in the HIIT group. Interestingly, the CFI was the only frailty tool to capture a statistically significant increase in frailty score in the SED mice between baseline and endpoint. There are several possibilities as to why this may be the case. First, the 8-week timeframe of this study may not have been sufficient to observe substantial natural decline in physical performance in these mice. Indeed, the C-FI uses a wider variety of parameters beyond physical performance which may have enable this tool to identify differences between 24 and 26 month old mice. Additionally, the C-FI criteria are derived from independent measurements of parameters, whereas the parameters from Fried based animal frailty tools are determined based upon relative differences within the cohort, which may obfuscate detection of frailty decline in control groups during short longitudinal studies. Finally, the Seldeen and Valencia score scales redefine cutoffs at each time point that could also diminish the contribution of aging to frailty assessment.

One possible limitation in this study was that we used cohort means to establish scores for the body weight and temperature parameters. Our decision to deviate from using reference cohort data as described in Whitehead et al. (2014)), was that we were concerned the diversity in body weight in our aged female mouse cohorts reduced the sensitivity of this parameter as most mice received a maximal score when our calculations were based on reference data. However, we do note as mentioned above that the C-FI used here successfully identified increase in frailty in the sedentary mice from baseline to endpoint. Additionally, another possible limitation is the use of transgenic mice in our study instead of true wild type mice. However, our anecdotal experience is that when the mutation is not induced in these mice that they exhibit similar physical performance and lifespan as true wild type mice. Furthermore, these mice also exhibit a similar decline in body weight typically observed in mice of advanced age (Anisimov et al., 2010; Dubrovsky et al., 2010), and that was also observed in our study of wild type male mice from 24 to 28 months of age obtained from the NIA (Seldeen et al., 2018a).

Both ADLs and IADLs have been surrogates for disability and functional capacity in clinical settings, and declines have been associated with both Fried based and deficit accumulation based frailty (Aguilar-Navarro et al., 2012; Kojima, 2018; Yang and Gu, 2016). Several spontaneous mouse behaviors are being examined as potential mouse equivalents of IADLs, including burrowing, scavenging, and nest building (Deacon, 2012). Nest building in particular is sometimes used as a measure of health in laboratory facilities (Jirkof, 2014) and has been shown to decline with aging (Xiong et al., 2018). Here we demonstrate that HIIT improves nest building ability in mice, which is consistent with human clinical studies that report that exercise is beneficial for ADLs and IADLs (Edgren et al., 2015; Nascimento et al., 2014; Osuka et al., 2018). We further determined that the Seldeen, Valencia Score, and FIAV frailty assessment tools correlate with nest building ability at baseline, although none correlated with nest building scores at end point. This may be due to a lack of sensitivity of

the categorical based scales. However, FIAV, with a continuous scoring system, was nearing correlation ( $p = 0.10$ ).

Most of the studied mouse frailty tools differentiated between surviving and non-surviving mice, a finding that is consistent with conclusions from studies involving the C-FI and Valencia score scales (Martinez de Toda et al., 2018; Kane et al., 2017b). However, none of the tools correlated with the number of days survived in the 7 deceased mice, with the FIAV tool being closest in this regard ( $p = 0.13$ ). This latter result suggests our study may not have been sufficiently powered with just 7 mice for this outcome, and a study with survival as the primary outcome is warranted. Despite the low power of this aspect of our study, we found two parameters to be associated with survival in these mice - body temperature and gait speed. Low body temperature is a well-established predictor of mortality in mice (Reynolds et al., 1985). Additionally gait speed has been shown to be a predictor of survival in human studies (Studenski et al., 2011). The frailty assessment scores did not appear to associate with underlying pathology, even though multiple human studies reveal that frailty is commonly associated with multiple co-morbidities. Our further analysis also revealed that none of the physical performance parameters correlated with underlying pathology. This suggests the mice can maintain physical function in light of the presence of tumors or organ dysfunction. We also note that although our investigators have extensive experience in working with mice none were formally trained in mouse pathology. Thus, this finding may indicate a need for greater sophistication in the analysis of pathology, including tissue histology, and calls for future studies to explore the relationship between underlying pathology and frailty status in mice.

Interestingly, our cohort exhibited great diversity in physical performance and body weight, which ranged from 21.0 g to 63.8 g at the start of the exercise phase. While epigenetics may have been a factor, our primary concern was the influence of animal cage conditions, as the mice were bred and housed in a ventilated caging system then moved at 21 months of age to wire top cages at the start of this experiment. There was considerable improvement in the SED mice after the transfer to the wire top cages, including loss of body weight, improved activity wheel usage, and improved open field activity. We speculate that the ventilated caging system may not provide as many climbing opportunities, as compared to the wire-top cages, which may cause obese mice to have less activity and exacerbate weight gain. Whether this caused the large diversity in body shape in our cohort is uncertain, as well as whether shifting to wire-top caging provided an environment that would foster improvement in physical performance. However, despite factors such as mouse body-type diversity, background, and environment, our data show that our HIIT protocol improved physical performance and thus enhances the clinical relevance to human studies that may also feature significant variation in body composition.

Our study to our knowledge represents the first to examine female mice with a Fried based animal frailty assessment tool – although we note that the C-FI has been used to examine female cohorts previously (Kane et al., 2016; Keller et al., 2018; Kane et al., 2019). In comparison to our study in male mice (Seldeen et al., 2018a) that used the Seldeen et al frailty tool, 1 frail and 8 prefrail mice were identified in a cohort of 24 mice at 24 months of age, and although this study did not identify any frail with this tool, it similarly identified

9 prefrail mice in a cohort of 25 female mice at 24-months of age. This count includes only mice that survived to the end of the study, as deceased mice were not included in the analysis of our previous study (Seldeen et al., 2018a). In contrast, female mice exhibited greater frailty in two of the studies using the C-FI at 17 and 22months of age (Keller et al., 2018; Kane et al., 2019), while the third study showed greater C-FI scores in male mice at 18months of age (Kane et al., 2016). Interestingly, Kane et al also reported that a frailty index based upon serum inflammatory biomarkers (called the FI-Lab) indicated male mice had greater frailty, and a combined index that pairs the FI-Lab with the C-FI found no difference between male and female mice (Kane et al., 2019). Along with our study these findings suggest that the parameters chosen may influence whether gender biases exist within a frailty assessment tool, and as suggested by Kane et al the inclusion of more parameters may dampen the effect of gender (Kane et al., 2019). This notion is also supported by a study by Kulminiski et al. that examined the morbidity-mortality paradox in humans that females have worse health but better survival than males. The authors found that an index of 34 minor health deficits exhibited no sex-specific patterns, while subsets that characterized male and female specific disadvantages identified opposite opposing and confirming support for the morbidity-mortality paradox, respectively (Kulminiski et al., 2008). These findings speak to the need for more studies investigating the impacts of gender on frailty assessment in mice, particularly for Fried based tools.

## 5. Conclusion

Short session high intensity interval training, administered 3-days-a-week, improves both physical performance and functional capacity in aged female mice. HIIT increases FIAV scores indicating lower frailty, and the observed trends across the Fried physical frailty based assessment tools further suggest HIIT reduces frailty. Additionally, most Fried based frailty assessment tools correlated to functional capacity scores and distinguished between surviving and non-surviving mice. Interestingly, the deficit accumulation tool was the only one to identify age-associated increase in frailty in SED mice. This study suggests Fried based and deficit accumulation based animal frailty tools may identify different aspects of frailty and that all of these instruments are valuable to lay the foundation for future studies that explore interventions for frailty, including HIIT, in clinically relevant animal models and ultimately human trials.

## Acknowledgements

The authors thank Jolie McCutcheon, Susan Campbell, Jasmine Lopez, and Deborah Heterbring, as well as other staff members of the University at Buffalo Laboratory Animal Facilities for husbandry, veterinary support, and care of the mice. The authors also wish to thank the University at Buffalo and the research service of the VA Western New York Healthcare System.

## Funding sources

This work was supported by the Veteran Affairs Rehabilitation Research and Development Service (Grant RX001066) and the Indian Trail Foundation.

## Abbreviations:

**ADLs**                      Activities of daily living

<b>C-FI</b>	Mouse clinical frailty index
<b>CXCL1</b>	Chemokine (C-X-C motif) ligand 1
<b>HIIT</b>	High intensity interval training
<b>IADLs</b>	Instrumental activities of daily living
<b>SED</b>	Sedentary

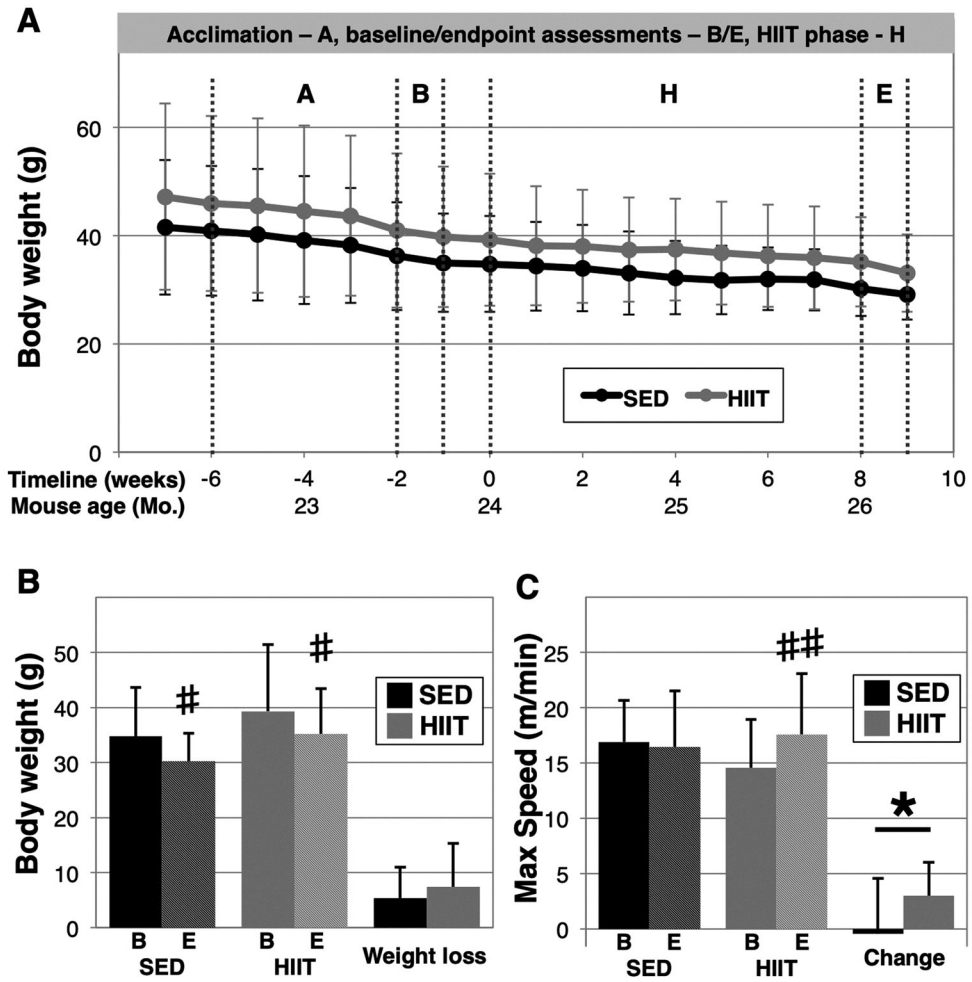
## References

- Aguilar-Navarro S, et al. , 2012. The phenotype of frailty predicts disability and mortality among mexican community-dwelling elderly. *J. Frailty Aging* 1 (3), 111–117. [PubMed: 27093198]
- Anisimov VN, et al. , 2010. Gender differences in metformin effect on aging, life span and spontaneous tumorigenesis in 129/Sv mice. *Aging (Albany NY)* 2 (12), 945–958. [PubMed: 21164223]
- Antoch MP, et al. , 2017. Physiological frailty index (PFI): quantitative in-life estimate of individual biological age in mice. *Aging (Albany NY)* 9 (3), 615–626. [PubMed: 28325885]
- Apostolo J, et al. , 2018. Effectiveness of interventions to prevent pre-frailty and frailty progression in older adults: a systematic review. *JBIC Database Syst. Rev. Implement. Rep* 16 (1), 140–232.
- Bell KE, et al. , 2015. Day-to-Day changes in muscle protein synthesis in recovery from resistance, aerobic, and high-intensity interval exercise in older men. *J. Gerontol. A Biol. Sci. Med. Sci* 70 (8), 1024–1029. [PubMed: 25650305]
- Bruseghini P, et al. , 2015. Effects of eight weeks of aerobic interval training and of isoinertial resistance training on risk factors of cardiometabolic diseases and exercise capacity in healthy elderly subjects. *Oncotarget* 6 (19), 16998–17015. [PubMed: 26046575]
- Cameron ID, et al. , 2013. A multifactorial interdisciplinary intervention reduces frailty in older people: randomized trial. *BMC Med* 11, 65. [PubMed: 23497404]
- Cheng HL, et al. , 2003. Developmental defects and p53 hyperacetylation in Sir2 homolog (SIRT1)-deficient mice. *Proc. Natl. Acad. Sci. U. S. A* 100 (19), 10794–10799. [PubMed: 12960381]
- Clark TC, et al., 2017. Early release of selected estimates based on data from 2016 National Health Interview Survey National Center for Health Statistics.
- Collard RM, et al. , 2012. Prevalence of frailty in community-dwelling older persons: a systematic review. *J. Am. Geriatr. Soc* 60 (8), 1487–1492. [PubMed: 22881367]
- Deacon R, 2012. Assessing burrowing, nest construction, and hoarding in mice. *J. Vis. Exp* 59, e2607.
- Dubrovsky YV, Samsa WE, Kondratov RV, 2010. Deficiency of circadian protein CLOCK reduces lifespan and increases age-related cataract development in mice. *Aging (Albany NY)* 2 (12), 936–944. [PubMed: 21149897]
- Edgren J, et al. , 2015. Effects of a home-based physical rehabilitation program on physical disability after hip fracture: a randomized controlled trial. *J. Am. Med. Dir. Assoc* 16 (4), 350 e1–7.
- Fiatarone MA, et al. , 1990. High-intensity strength training in nonagenarians. Effects on skeletal muscle. *JAMA* 263 (22), 3029–3034. [PubMed: 2342214]
- Fried LP, et al. , 2001. Frailty in older adults: evidence for a phenotype. *J. Gerontol. A Biol. Sci. Med. Sci* 56 (3), M146–M156. [PubMed: 11253156]
- Gaskill BN, et al. , 2013. Nest building as an indicator of health and welfare in laboratory mice. *J. Vis. Exp* 82, 51012.
- Gomez-Cabrera MC, et al. , 2017. A new frailty score for experimental animals based on the clinical phenotype: inactivity as a model of frailty. *J. Gerontol. A Biol. Sci. Med. Sci* 72 (7), 885–891. [PubMed: 28329258]
- Gonzalez-Vaca J, et al. , 2014. Frailty in INstitutionalized older adults from ALbacete. The FINAL Study: rationale, design, methodology, prevalence and attributes. *Maturitas* 77 (1), 78–84. [PubMed: 24189222]

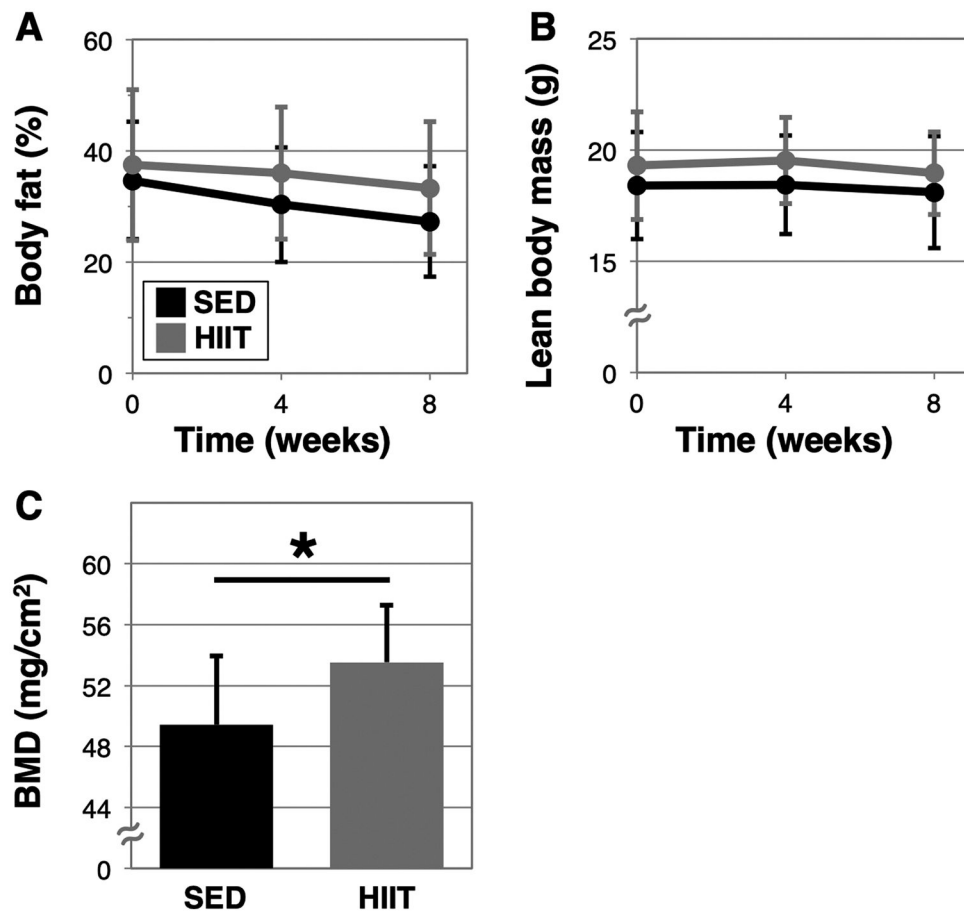
- Graber TG, et al. , 2015. Voluntary aerobic exercise reverses frailty in old mice. *J. Gerontol. A Biol. Sci. Med. Sci* 70 (9), 1045–1058. [PubMed: 25271307]
- Gulley SP, et al. , 2018. At the intersection of chronic disease, disability and health services research: a scoping literature review. *Disabil. Health J* 11 (2), 192–203. [PubMed: 29396271]
- Jirkof P, 2014. Burrowing and nest building behavior as indicators of well-being in mice. *J. Neurosci. Methods* 234, 139–146. [PubMed: 24525328]
- Justice JN, et al. , 2014. Battery of behavioral tests in mice that models age-associated changes in human motor function. *Age (Dordr)* 36 (2), 583–592. [PubMed: 24122289]
- Justine M, et al. , 2013. Barriers to participation in physical activity and exercise among middle-aged and elderly individuals. *Singapore Med. J* 54 (10), 581–586. [PubMed: 24154584]
- Kane AE, et al. , 2016. Impact of longevity interventions on a validated mouse clinical frailty index. *J. Gerontol. A Biol. Sci. Med. Sci* 71 (3), 333–339. [PubMed: 25711530]
- Kane AE, et al. , 2017a. A comparison of two mouse frailty assessment tools. *J. Gerontol. A Biol. Sci. Med. Sci* 72 (7), 904–909. [PubMed: 28549083]
- Kane AE, et al. , 2017b. The association between frailty, the metabolic syndrome, and mortality over the lifespan. *Geroscience* 39 (2), 221–229. [PubMed: 28281219]
- Kane AE, et al. , 2019. A murine frailty index based on clinical and laboratory measurements: links between frailty and pro-inflammatory cytokines differ in a sex-specific manner. *J. Gerontol. A Biol. Sci. Med. Sci* 74 (3), 275–282. [PubMed: 29788087]
- Keller K, et al. , 2018. Chronic treatment with the ACE inhibitor enalapril attenuates the development of frailty and differentially modifies pro-and anti-inflammatory cytokines in aging male and female C57BL/6 mice. *J. Gerontol. A Biol. Sci. Med. Sci* E-published
- Knowles AM, et al. , 2015. Impact of low-volume, high-intensity interval training on maximal aerobic capacity, health-related quality of life and motivation to exercise in ageing men. *Age (Dordr)* 37 (2), 25. [PubMed: 25773069]
- Kojima G, 2018. Quick and simple FRAIL scale predicts incident activities of daily living (ADL) and instrumental ADL (IADL) disabilities: a systematic review and meta-analysis. *J. Am. Med. Dir. Assoc* 19 (12), 1063–1068. [PubMed: 30206033]
- Kulminski AM, et al. , 2008. Sex-specific health deterioration and mortality: the morbidity-mortality paradox over age and time. *Exp. Gerontol* 43 (12), 1052–1057. [PubMed: 18835429]
- Li G, et al. , 2015. Comparison between frailty index of deficit accumulation and phenotypic model to predict risk of falls: data from the global longitudinal study of osteoporosis in women (GLOW) hamilton cohort. *PLoS One* 10 (3), e0120144. [PubMed: 25764521]
- Liu H, et al. , 2014. Clinically relevant frailty index for mice. *J. Gerontol. A Biol. Sci. Med. Sci* 69 (12), 1485–1491. [PubMed: 24336799]
- Losa-Reyna J, et al. , 2019. Effect of a short multicomponent exercise intervention focused on muscle power in prefrail and frail elderly: a pilot trial. *Exp. Gerontol* E-published. pubmed id: 30528641.
- Manas A, et al. , 2018. Reallocating accelerometer-assessed sedentary time to light or moderate-to vigorous-intensity physical activity reduces frailty levels in older adults: an isothermal substitution approach in the TSHA study. *J. Am. Med. Dir. Assoc* 19 (2), 185 e1–185 e6.
- Martinez de Toda I, et al. , 2018. Frailty quantified by the “Valencia score” as a potential predictor of lifespan in mice. *J. Gerontol. A Biol. Sci. Med. Sci* 73 (10), 1323–1329. [PubMed: 29718119]
- Nascimento CM, et al. , 2014. Effect of a multimodal exercise program on sleep disturbances and instrumental activities of daily living performance on Parkinson’s and Alzheimer’s disease patients. *Geriatr. Gerontol. Int* 14 (2), 259–266. [PubMed: 23647635]
- Osuka Y, et al. , 2018. Association between exercise type and the decline in instrumental activities of daily living in community-dwelling older women: a 4-year prospective study. *Prev. Med* 112, 23–30. [PubMed: 29596918]
- Parks RJ, et al. , 2012. A procedure for creating a frailty index based on deficit accumulation in aging mice. *J. Gerontol. A Biol. Sci. Med. Sci* 67 (3), 217–227. [PubMed: 22021390]
- Reynolds MA, Ingram DK, Talan M, 1985. Relationship of body temperature stability to mortality in aging mice. *Mech. Ageing Dev* 30 (2), 143–152. [PubMed: 4021553]



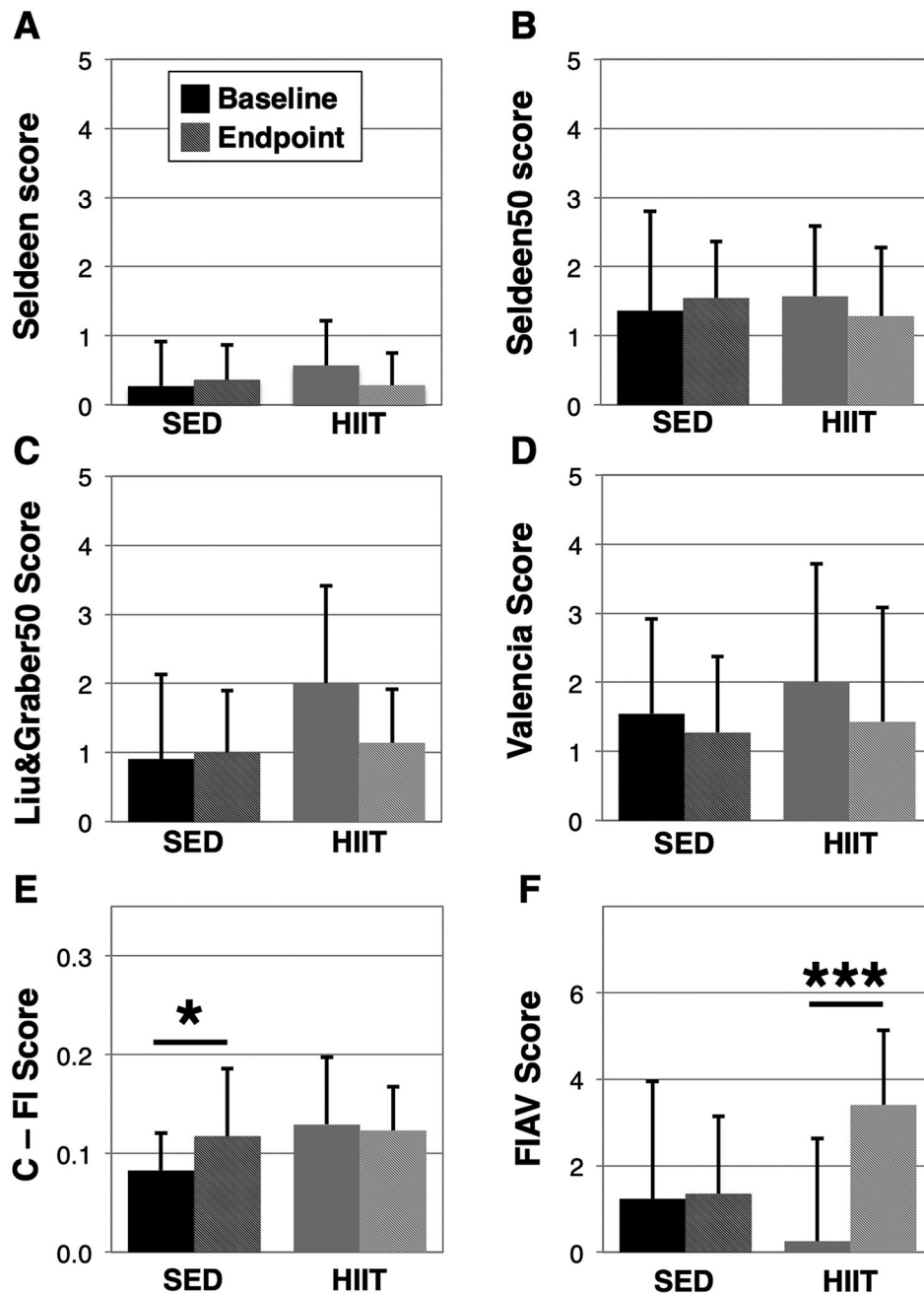
- Rockwood K, et al. , 2005. A global clinical measure of fitness and frailty in elderly people. *CMAJ* 173 (5), 489–495. [PubMed: 16129869]
- Rogers NT, et al. , 2017. Physical activity and trajectories of frailty among older adults: evidence from the English Longitudinal Study of Ageing. *PLoS One* 12 (2), e0170878. [PubMed: 28152084]
- Schutzer KA, Graves BS, 2004. Barriers and motivations to exercise in older adults. *Prev. Med* 39 (5), 1056–1061. [PubMed: 15475041]
- Sculthorpe N, Herbert P, Grace FM, 2015. Low-frequency high-intensity interval training is an effective method to improve muscle power in lifelong sedentary aging men: a randomized controlled trial. *J. Am. Geriatr. Soc* 63 (11), 2412–2413. [PubMed: 26603065]
- Seldeen KL, et al. , 2018a. High intensity interval training improves physical performance and frailty in aged mice. *J. Gerontol. A Biol. Sci. Med. Sci* 73 (4), 429–437. [PubMed: 28633487]
- Seldeen KL, Redae Y, Thiyagarajan R, Berman RN, Leiker MM, Troen BR, 2018b. Short session high intensity interval training and treadmill assessment in aged mice. *J. Vis. Exp.*, e59138.
- Studenski S, et al. , 2011. Gait speed and survival in older adults. *JAMA* 305 (1), 50–58. [PubMed: 21205966]
- Whitehead JC, et al. , 2014. A clinical frailty index in aging mice: comparisons with frailty index data in humans. *J. Gerontol. A Biol. Sci. Med. Sci* 69 (6), 621–632. [PubMed: 24051346]
- Xiong XD, et al. , 2018. Age- and gender-based differences in nest-building behavior and learning and memory performance measured using a radial six-armed water maze in C57BL/6 mice. *Behav. Neurol* 2018 8728415. [PubMed: 29854021]
- Yamada M, et al. , 2012. Community-based exercise program is cost-effective by preventing care and disability in Japanese frail older adults. *J. Am. Med. Dir. Assoc* 13 (6), 507–511. [PubMed: 22572554]
- Yang F, Gu D, 2016. Predictability of frailty index and its components on mortality in older adults in China. *BMC Geriatr* 16, 145. [PubMed: 27455962]



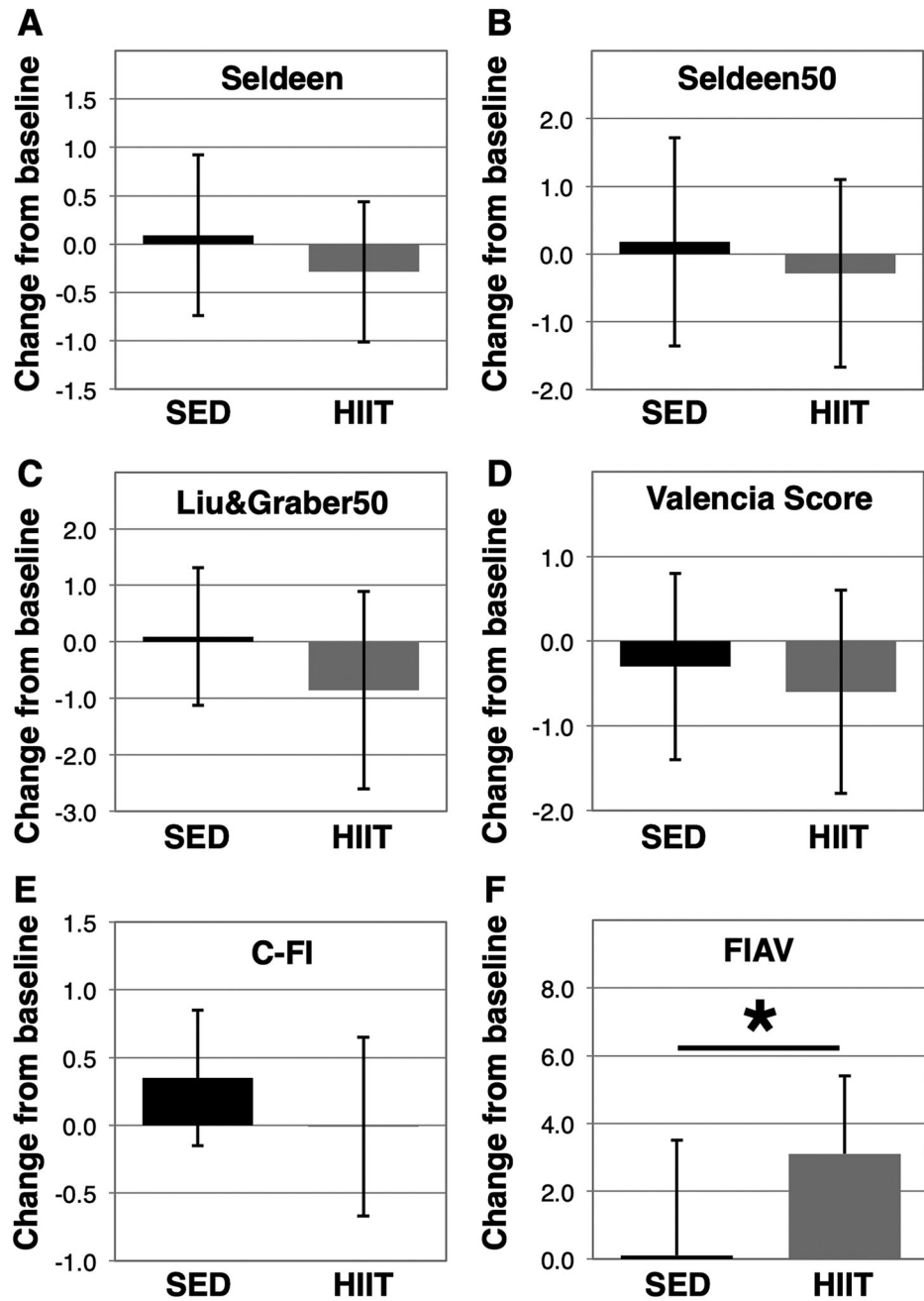
**Fig. 1.** HIIT improves physical performance in aged female mice. The experimental timeline begins at 21-months of age in female C57BL/6 J mice (A). At 24-months of age, mice were administered HIIT ( $n = 14$ ) or remained sedentary (SED,  $n = 11$ ) for 8 weeks. Body weight was measured weekly (A) and the change in body weight was assessed from the start to the end of the HIIT phase (B). Baseline uphill interval treadmill (C) was used to assign the initial intensity for HIIT exercise and as a marker for subsequent impacts of the exercise. A “\*” indicates  $p < 0.05$  in an unpaired  $t$ -test between groups, a “#” and a “##” indicates  $p < 0.05$  and  $p < 0.01$ , respectively in a paired  $t$ -test within groups, “B” indicates baseline, and “E” indicates endpoint.



**Fig. 2.** Changes in body composition and bone density in response to HIIT. Body fat percentage (A) and lean body mass (B) were measured in mice at baseline, after 4-weeks, and after 8-weeks (endpoint) of HIIT ( $n = 14$ ) or remaining sedentary (SED,  $n = 11$ ), using quantitative magnetic resonance (qMR). At endpoint, bone mineral density (BMD) was assessed using DEXA (C), SED  $n = 7$ , HIIT  $n = 11$ , \* indicates  $p < 0.05$ .

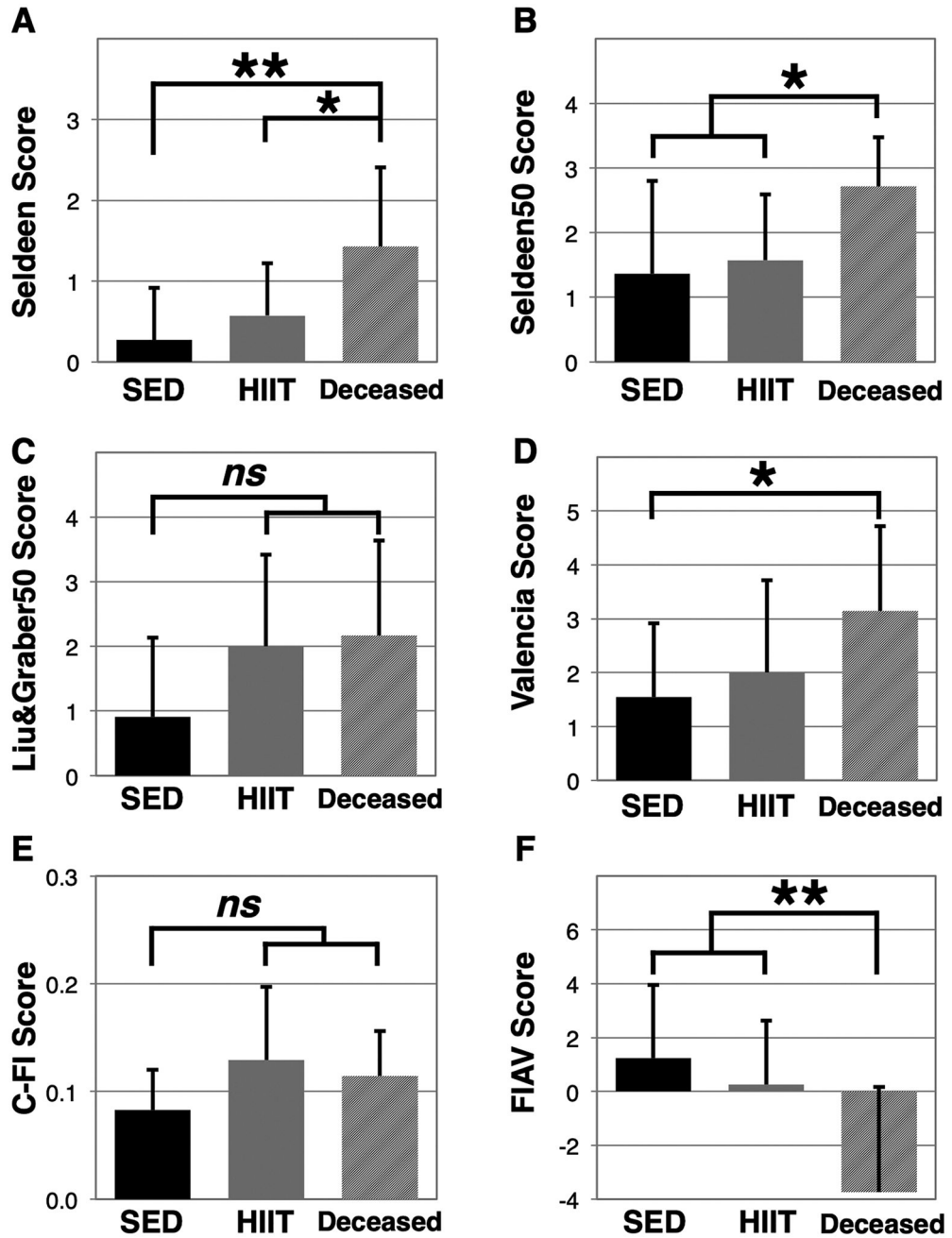


**Fig. 3.** Assessment of the impacts of HIIT on frailty using multiple frailty assessment tools. The frailty status of SED and HIIT mice was assessed before and after the 8-week exercise phase using the Seldeen (A), Seldeen50 (B), Liu&Graber50 (C), Valencia score (D), C-FI (E), and FIAV (F) animal frailty assessment tools. Darker colors represent baseline and lighter colors represent endpoint for SED and HIIT, respectively. A “\*” and “\*\*\*” indicates a  $p < 0.05$  and  $p < 0.001$  in a paired  $t$ -test within groups, respectively. Higher scores are indicative of greater frailty in all scales except for FIAV (F).

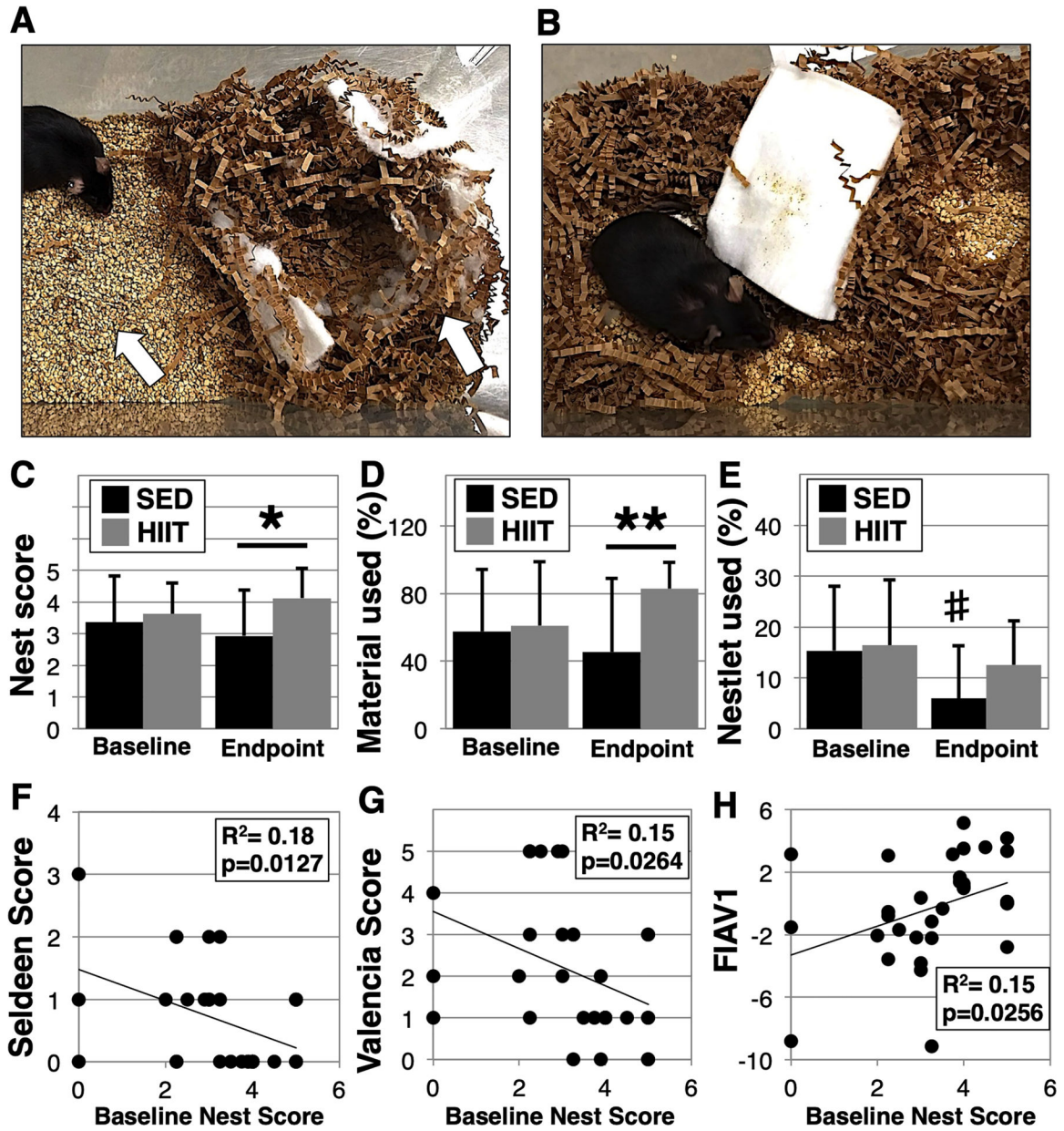


**Fig. 4.**

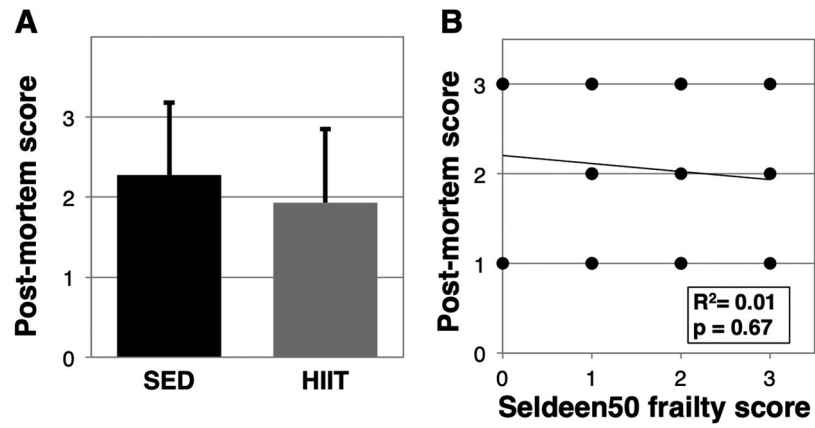
Change in frailty scores from baseline to endpoint. Frailty status was assessed in 24-month female mice and again 8-weeks later after administration of a HIIT exercise regimen (HIIT,  $n = 14$ ) or remaining sedentary (SED,  $n = 11$ ), and the change in frailty score was determined using the Seldeen (A), Seldeen50 (B), Liu&Graber50 (C), Valencia score (D), C-FI (E), and FIAV (F) animal frailty assessment tools. A “\*” indicates  $p < 0.05$  in HIIT versus SED. Higher scores are indicative of increased frailty in all scales except for FIAV.



**Fig. 5.** Frailty assessment tools distinguish between surviving and non-surviving mice. Baseline frailty status was compared between sedentary mice (SED,  $n = 11$ ), HIIT mice ( $n = 14$ ), and mice that were deceased prior to endpoint assessments (deceased,  $n = 7$ , except Liu&Graber50,  $n = 5$ ) using the Seldeen (A), Seldeen50 (B), Liu&Graber50 (C), Valencia score (D), C-FI (E), and FIAV (F) animal frailty assessment tools. A “\*” and “\*\*” indicates  $p < 0.05$  and  $p < 0.01$ , respectively, in comparison between SED and deceased mice. “*ns*” denotes the comparison was not significant. Higher scores are indicative of greater frailty in all scales except for FIAV.



**Fig. 6.** HIIT improves nest building ability. Nest building in SED and HIIT mice was determined at baseline and endpoint assessing the quality of nest, the amount of material cleared and put into the nest (A, left arrow), and the incorporation of a cotton pad nestlet (A, right arrow). Panel B shows a representative mouse with a poorly constructed nest and no use of nesting or nestlet materials. Nests were scored on a range of 0-5 (C), while paper nesting material used (B) and cotton nestlet material used (E) were estimated as a percent of total. Next, baseline frailty scores were correlated against baseline nest scores and three tools were found to have significance ( $p < 0.05$ ) including Seldeen (F), Valencia (G), and FIAV (H). A “\*” and “\*\*\*” indicates a  $p < 0.05$  and  $p < 0.01$  in an unpaired  $t$ -test between groups, respectively. A “#” indicates  $p < 0.05$  in a paired  $t$ -test within a group.



**Fig. 7.** Frailty tools fail to predict underlying pathology. Following endpoint assessments, the mice were culled and a post-mortem inspection was performed. To estimate the condition, mice with identified tumors, enlarged organs, or other serious abnormalities were scored a 1, minor abnormalities a 2, and otherwise unremarkable a 3. The scoring system was first applied to compare SED versus HIIT mice (A). Next, the score system was correlated to each of the frailty assessment tools, however, (B) shows a representative comparison as correlation was not observed with any of the frailty tools.



Physical performance assessments in SED and HIIT mice. 24-month female mice were assessed using a battery of physical performance behavioral tests and then were administered an 8-week HIIT regimen ( $n = 14$ ) or remained sedentary ( $n = 11$ ) for 8 weeks. Baseline and endpoint performance was then compared within groups using a paired student's  $t$ -test, while the changes from baseline to endpoint were compared using an unpaired student's  $T$ -test ( $p < 0.05$  indicated in bold). We did not identify any significant differences between groups at either time point.

Table 1

Assessment	SED			HIIT			
	Baseline	Endpoint	p	Baseline	Endpoint	p	
Uphill treadmill (m/min)	16.9 ± 3.8	16.5 ± 5.0	0.7713	14.6 ± 4.3	17.6 ± 5.5	<b>0.0025</b>	SED Delta: -0.5 ± 5.0 HIIT Delta: 3.0 ± 3.0 <b>0.0441</b>
Flat treadmill (min)	19.9 ± 10.0	25.3 ± 7.3	0.1510	18.2 ± 10.5	31.8 ± 13.7	< <b>0.0001</b>	5.3 ± 11.4 13.6 ± 7.5 <b>0.0114</b>
Grip strength (N)	2.07 ± 0.37	1.73 ± 0.19	<b>0.0037</b>	2.00 ± 0.31	1.87 ± 0.30	0.2534	-0.34 ± 0.30 -0.13 ± 0.40 0.1636
Grip str/BW (mN/g)	59.8 ± 13.3	51.1 ± 13.4	<b>0.0031</b>	54.4 ± 20.6	49.1 ± 11.8	0.1812	-8.8 ± 7.4 -5.4 ± 14.2 0.4869
Gait speed (cm/s)	29.8 ± 6.2	32.0 ± 6.1	0.2604	28.0 ± 5.2	32.7 ± 5.5	<b>0.0280</b>	2.2 ± 5.9 4.7 ± 7.0 0.3409
Rotarod latency (s)	105.8 ± 39.8	125.2 ± 46.5	0.0892	103.8 ± 65.5	122.1 ± 61.7	0.3601	19.4 ± 34.1 18.3 ± 72.2 0.9648
Grid hang latency (s)	27.1 ± 35.4	32.6 ± 43.3	0.2711	38.6 ± 72.9	46.2 ± 85.1	0.2640	5.5 ± 15.8 7.6 ± 24.5 0.8078
Grid hang impulse (N.s)	8.1 ± 8.7	8.5 ± 9.5	0.8177	10.7 ± 17.3	12.1 ± 20.6	0.4705	0.4 ± 5.6 1.4 ± 7.0 0.7050
Activity wheels (km/d)	0.33 ± 0.21	0.47 ± 0.17	<b>0.0006</b>	0.24 ± 0.24	0.48 ± 0.33	<b>0.0023</b>	0.15 ± 0.24 0.24 ± 0.27 0.4255
Open field dis. (km)	12.1 ± 8.3	22.7 ± 9.7	<b>0.0004</b>	12.7 ± 8.6	24.4 ± 10.0	<b>0.0022</b>	10.5 ± 6.1 11.7 ± 11.6 0.7698
OF crossings	64.1 ± 51.7	104.2 ± 45.9	<b>0.0046</b>	61.2 ± 44.3	114.0 ± 46.7	<b>0.0050</b>	40.1 ± 33.9 52.8 ± 58.5 0.5456
OF rearings	39.9 ± 28.9	79.2 ± 50.5	<b>0.0006</b>	50.2 ± 67.3	73.9 ± 50.5	0.1009	39.3 ± 24.1 23.6 ± 50.1 0.3719
OF velocity (cm/s)	11.1 ± 1.3	12.8 ± 1.0	<b>0.0010</b>	11.1 ± 1.6	12.1 ± 1.4	<b>0.0283</b>	1.6 ± 1.1 0.9 ± 1.4 0.2188

Correlation between Frailty assessment tools. The correlation coefficient ( $r$ ) was determined, using both SED and HIIT mice ( $N = 25$ ), between each frailty assessment tool at baseline and at endpoint.

**Table 2**

<b>Baseline</b>					
	<b>Seldeen50</b>	<b>Liu&amp;Graber50</b>	<b>Valencia</b>	<b>C-FI</b>	<b>FLAV</b>
Seldeen	0.68 *	0.60 *	0.67 *	0.42 *	-0.65 *
Seldeen50		0.48 *	0.55 *	0.32	-0.68 *
Liu&Graber50			0.77 *	0.29	-0.70 *
Valencia				0.09	-0.52 *
C-FI					-0.40
<b>Endpoint</b>					
	<b>Seldeen50</b>	<b>Liu&amp;Graber50</b>	<b>Valencia</b>	<b>C-FI</b>	<b>FLAV</b>
Seldeen	0.17	-0.34	-0.36	0.13	-0.17
Seldeen50		-0.02	0.34	0.01	-0.57 *
Liu&Graber50			0.54 *	-0.02	-0.07
Valencia				0.03	-0.36
C-FI					0.18

An \* indicates  $p < 0.05$ .

Comparison of physical performance in surviving versus non-surviving mice. Baseline parameters were compared in SED ( $n = 11$ ) and mice that died before endpoint assessments ( $n=7$ ). Additionally, assessment values were correlated to the number of days survived from baseline in the group of deceased mice.

Table 3

	SED	Deceased	p	Correlation (r) to days survived
Body weight (g)	39.2 ± 11.8	44.8 ± 20.3	0.4657	0.67
Weight loss 1 wk. prior	-0.5 ± 2.4	-5.3 ± 7.5	0.0667	0.68
Weight loss 1 mo. prior	-9.9 ± 6.2	-14.9 ± 8.0	0.1496	-0.01
Body temperature (°C)	32.3 ± 0.6	32.1 ± 1.0	0.6536	<b>-0.84*</b>
Uphill treadmill (m/min)	14.9 ± 3.5	13.4 ± 5.8	0.5060	-0.21
Flat treadmill (min)	18.1 ± 8.1	14.2 ± 11.9	0.4211	0.24
Grip strength (N)	2.10 ± 0.40	1.80 ± 0.30	0.1306	0.41
Gait speed (cm/s)	31.0 ± 5.9	22.7 ± 6.8	<b>0.0260*</b>	-0.66
Rotarod latency (s)	105.8 ± 39.8	80.8 ± 76.8	0.3828	-0.39
Grip grid latency (s)	32.7 ± 34.9	79.7 ± 101.7	0.1763	-0.18
Rotarod + Grip grid (s)	138.5 ± 61.4	160.4 ± 171.4	0.7027	-0.63
Activity wheels (km/d)	0.32 ± 0.21	0.19 ± 0.33	0.3600	-0.46
Open field distance (m)	12.6 ± 8.1	9.8 ± 5.7	0.4299	0.20
Open field crossings	65.6 ± 49.3	47.3 ± 29.7	0.3908	0.22
Open field rearings	39.8 ± 27.4	30.0 ± 27.2	0.4683	0.33
Open field velocity (cm/s)	11.1 ± 1.2	10.7 ± 2.1	0.5717	0.04

A \* indicates  $p < 0.05$ .