OXFORD

THE GERONTOLOGICAL SOCIETY OF AMERICA®

Original Article

Measuring Exercise Capacity and Physical Function in Adult and Older Mice

Ted G. Graber, PhD,[1,][*,] **Rosario Maroto, MD, PhD,**[2] **Christopher S. Fry, PhD,**[2,3] **Camille R. Brightwell, MS,**[3,4] **and Blake B. Rasmussen, PhD**[2,]

[1]Department of Physical Therapy, East Carolina University, Greenville, North Carolina. [2]Department of Nutrition and Metabolism, University of Texas Medical Branch, Galveston. [3]Present address: Department of Athletic Training and Clinical Nutrition, University of Kentucky, Lexington. [4]Cell Biology PhD Program, University of Texas Medical Branch, Galveston.

*Address correspondence to: Ted G. Graber, PhD, Department of Physical Therapy, East Carolina University, 600 Moye Blvd, HSB Rm 2410, Mail Stop 668, Greenville, NC 27858-4353. E-mail: grabert19@ecu.edu

Abstract

The inability of older adults to maintain independence is a consequence of sarcopenia and frailty. In order to identify the molecular mechanisms responsible for decreased physical function, it will be critical to utilize a small animal model. The main purpose of this study was to develop a composite Comprehensive Functional Assessment Battery (CFAB) of well-validated tests to determine physical function and exercise capacity in 3 age groups of male C57BL/6 mice (6 months old, *n* = 29; 24 months old, *n* = 24; 28+ months old, *n* = 28). To measure physical function in mice, we used rotarod (overall motor function), grip meter (forelimb strength), treadmill (endurance), inverted cling (strength/endurance), voluntary wheel running (volitional exercise and activity rate), and muscle performance with in vivo contractile physiology (dorsiflexor torque). We hypothesized that CFAB would be a valid means to assess the physical function of a given mouse across the life span. In addition, we proposed that CFAB could be used to determine relationships between different parameters associated with sarcopenia. We found that there was an overall age-related significant decline (*p* < .05) in all measurements, and the CFAB score demonstrated that some individual mice (the upper quartile) retained the functional capacity of average mice 1 cohort younger. We conclude that the CFAB is a powerful, repeatable, and noninvasive tool to assess and compare physical function and assess complex motor task ability in mice, which will enable researchers to easily track performance at the individual mouse level.

Keywords: Frailty, Function, Mouse models, Muscle, Neuromuscular decline, Sarcopenia

The loss of physical function and exercise capacity that accompanies advanced age is highly significant, because many quality of life and health issues vitally depend upon maintaining mobility, strength, and endurance (1–3). Neuromuscular functional decline predates diseases such as sarcopenia and frailty, but there is potential for slowing the rate of decline using various modes of exercise (4–6). However, often exercise intolerance or disability can partly restrict exercise participation in exactly the population that needs it the most. Therefore, we need to examine the molecular mechanisms underlying both how age negatively affects the neuromuscular system and how therapies attenuate this loss. Animal models of exercise and functional ability are critical to developing novel therapeutic targets.

Our primary purpose was to validate a composite scoring system, using a repeatable, noninvasive methodology, that would enable a single number to describe the physical function capability of a given

mouse. This scoring system, Comprehensive Functional Assessment Battery (CFAB), enables powerful comparisons, within and between groups, of functional status at the individual level. This system is an evolutionary successor to our earlier work (5,7). Our secondary purpose was to evaluate how the CFAB determinants were associated with one another and with other measurements of age-related muscular decline (loss of muscle mass and strength). We hypothesized that some mice at older ages would retain relatively high levels of function similar to mice at younger ages. We further hypothesized that the CFAB system would prove to be a valuable resource by which to improve statistical power to detect differences compared to the individual tests alone.

To accomplish our primary aim to produce CFAB, we used validated noninvasive and repeatable determinants: treadmill, voluntary wheel running (VWR), grip strength, inverted cling, and rotarod (5,7–10). We tested 3 age groups of male C57BL/6 mice (6, 24,

and 28+ months of age). We then converted the individual determinants to the composite score CFAB. Overall, there was a significant age-related functional decline ($p < .05$) in the individual tests. The CFAB score demonstrated declining function over the life span, with the relative preservation of the upper quartile in both older groups. In addition, as is typical with sarcopenia, the older mice lost muscle mass and strength (measured with in vitro/ex vivo and in vivo contractile physiology) We concluded that CFAB was an excellent tool to improve the power of detection and will enable future work to more readily evaluate therapeutic efficacy regarding physical function status.

Method

Mice

C57BL/6 male mice of 3 different ages ($n = 29$, $n = 24$, and $n = 28$ at 6, 24, and 28+ months of age, respectively) were obtained from the National Institutes of Health/National Institute on Aging Aging Rodent Colony. Mice were treated humanely under approved Institutional Animal Care and Use Committee (IACUC) protocols. The mice were group-housed in 12-hour:12-hour light/dark cycle at 22°C and fed/watered ad libitum.

Functional Tests

For CFAB determination, we used *Treadmill (endurance capacity), Grip Test (forelimb strength), Inverted Cling (four limb strength/endurance), Rotarod (balance, coordination, gait speed, power generation)*, and *VWR (volitional exercise capacity and activity rate)*. A brief description of each procedure is given in Supplementary Material and details were previously published (5,7–10).

Comprehensive Functional Assessment Battery

Similar to previously described (5), the data from individual functional tests (rotarod, grip test, treadmill, inverted cling, and VWR) were converted into the CFAB composite score. Using the 6-month-old mice as the reference group, the test values of the 3 age groups were standardized (distance of each individual mouse's score from the mean 6-month value in units of the 6-month group standard deviation, *SD*). The standardized scores were summed into the single CFAB score, (units of total *SD*) representing a single number encompassing total physical functional ability. The inverted cling test data violated some assumptions of normality (Supplementary Table 1) and were transformed using log10.

Other Methods (Not Part of CFAB, Used to Document Sarcopenia)

In vitro and in vivo muscle contractile physiology
A subset of mice was tested for dorsiflexor (primarily TA or tibialis anterior) torque output (using in vivo contractile physiology) and the force/specific force (in vitro contractile physiology) of the EDL (extensor digitorum longus) and soleus (SOL) muscles of the mice. These test methods have been previously published (5–9). Physiological cross-sectional area (PCSA) was derived as previously established (5). See Supplementary Material for further details about the methodology.

Body and muscle mass
The mice were weighed to determine body mass during the CFAB procedure. At the time of tissue collection, muscles were blotted dry and weighed. We use a total muscle mass (sum of the muscle masses) to correlate physical function with muscle size. These measurements *were not* included in the CFAB score but are presented to help document sarcopenia.

Statistics

We report the mean ± standard error (SE) of the mean, unless otherwise noted. Statistical significance was $p < .05$, and trends reported where $.05 < p < .10$. Skew, kurtosis, and the Shapiro–Wilk test for the CFAB determinants are given in Supplementary Table 1, with few determinants (inverted cling being an exception) exhibiting excessive deviation from normality (skew <2.0 in most cases, and Shapiro–Wilk test is negative) (11–13). We transformed the inverted cling data using a log10 transformation, using the transformed values in CFAB. Analysis of covariance, adjusted for body mass, detected mean differences in CFAB components, and one-way analysis of variance for ex vivo and in vivo physiology, and mass. Post hoc analysis showed the least significant differences. In regression analysis, we report correlations as strong if $R > 0.7$, moderate if $0.50 < R < 0.75$, slight if $0.25 < R < 0.50$, and no relationship if $R < 0.25$. Effect sizes (14,15) and power analyses are given in Supplementary Material. Statistical analysis was performed using SPSS v24 and v26 (IBM) and GPower (16,17).

Results

Physical Function Tests That Comprise CFAB

Rotarod (balance, coordination, gait speed, and power generation)
The 6-month-old mice mean was 124.0 ± 5.3 seconds latency to fall, with 24-months-old (88.4 ± 6.1 s) 28% lower than the 6-months-old ($p < .001$). The 28+-months-old mice (76.4 ± 3.6) ran 39% less than 6-months-old mice ($p < .001$) and tended to run 14% less than 24-months-old ($p = .092$; Figure 1A).

Grip test (forelimb strength)
The 6-month-old mice mean was 1.4 ± 0.1 *N*, with 24-months-old (1.1 ± 0.2 *N*) 20% weaker than 6 months old ($p < .001$). The 28+-month-old mice (1.1 ± 0.0 *N*) were 25% weaker than 6-months-old ($p < .001$), but similar to 24-months-old (–7% numerically, but $p = .163$; Figure 1B).

Treadmill (endurance capacity)
The 6-month-old mice mean was 319.8 ± 23.7 seconds, with 24-month-old (350.7 ± 25.9 s) and 6-month-old mice having no significant difference (10% numeric difference, but $p = .382$). However, 28+-month-old mice (232.3 ± 25.1 s) were 27% lower than 6-month-old ($p = .014$) and 34% lower than 24-month-old mice ($p = .003$; Figure 1C).

Inverted cling (four limb strength/endurance)
The mean of the 6-month group was 183.17 ± 17.86 seconds latency to fall, with 24-month-old mice (115.0 ± 14.6 s) 37% lower ($p = .006$; Figure 1D). The 28+-month-old mice (87.3 ± 17.4 s) held on 52% less than 6-month-old mice ($p < .001$), but did not significantly differ from 24-month-old mice (numerically –24%, but nonsignificant at $p = .285$). Because the inverted cling test data violated some assumptions of normality (Supplementary Table 1), the
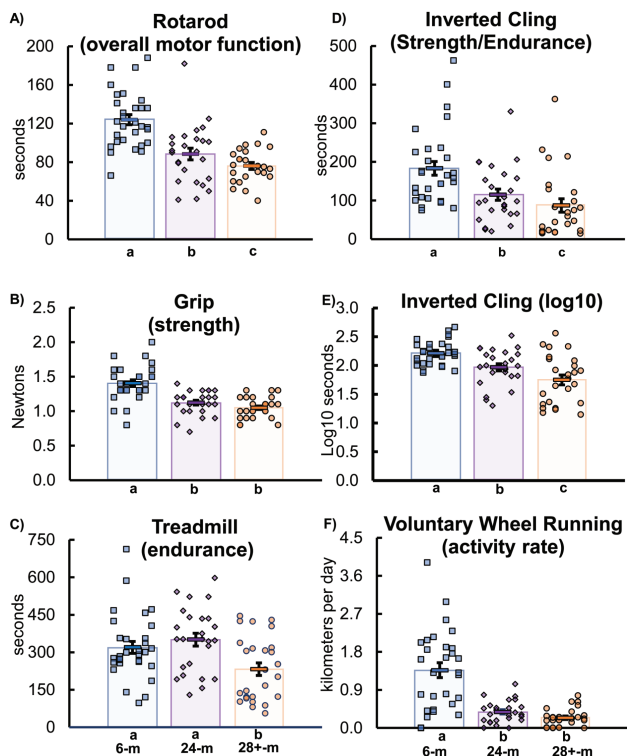
Figure 1. Functional tests: (**A**) rotarod performance, (**B**) grip strength, (**C**) treadmill, (**D**) inverted cling, (**E**) inverted cling (log10 transformation), and (**F**) voluntary wheel running. Key: Different letters indicate a significant difference ($p < .05$) between groups (from analysis of covariance adjusted for body mass and using least significant difference post hoc test). Each symbol (squares for mice 6 months old, diamonds for 24 months old, and circles for 28+ months old) indicates the test score for an individual mouse. The rectangles with error symbols (±SEM) in each grouping of ages indicate the mean value for that group. m = months; SEM = standard error of measurement. Full color version is available within the online issue.

log10 transformation was used to compare the data (Figure 1E) and to construct CFAB. The log10 6-month-old mice mean was 1.97 ± 0.06 log10 seconds latency to fall. The log10 24-month-old mice (1.97 ± 0.06) was 11% lower ($p = .001$) than the 6-month-old mice. The 28+-month-old mice (1.75 ± 0.09 log10 s) held on for a 20.9% shorter time than 6-month-old mice ($p < .001$) and were 11.1% lower ($p = .043$) than 24-month-old mice, on the log10 scale.

VWR (activity rate and volitional exercise capacity)

The 6-month-old mice mean was 1.37 ± 0.18 km/day. The 24-month-old mice (0.38 ± 0.05 km/day) ran 72% less than the 6-month-old mice ($p < .001$). The 28+-month-old mice (0.25 ± 0.05 km/day) tended to run 34% less than 24-month-old mice ($p = .080$) and 82% less distance than the 6-month-old mice ($p < .001$; Figure 1F).

Comprehensive Functional Assessment Battery

The composite CFAB score comprises the sum of the standardized scores (6-month group as the reference) from the rotarod, grip test, treadmill, inverted cling (log10 transformation applied), and VWR functional tests. In Figure 2, the 24-month-old mice ($n = 24$; −3.84 ± 0.11) were below ($p < .001$) the 6-month old ($n = 29$) mean CFAB score (−0.06 ± 0.44). The 28+-month-old ($n = 24$) mice (−5.74 ± 0.35) were below both the mean 6-month-old mice score ($p < .001$) and the 24-month-old mice ($p = .002$). Supplementary Figure 1A demonstrates that the 24-month upper quartile (top 25%) is functionally similar to
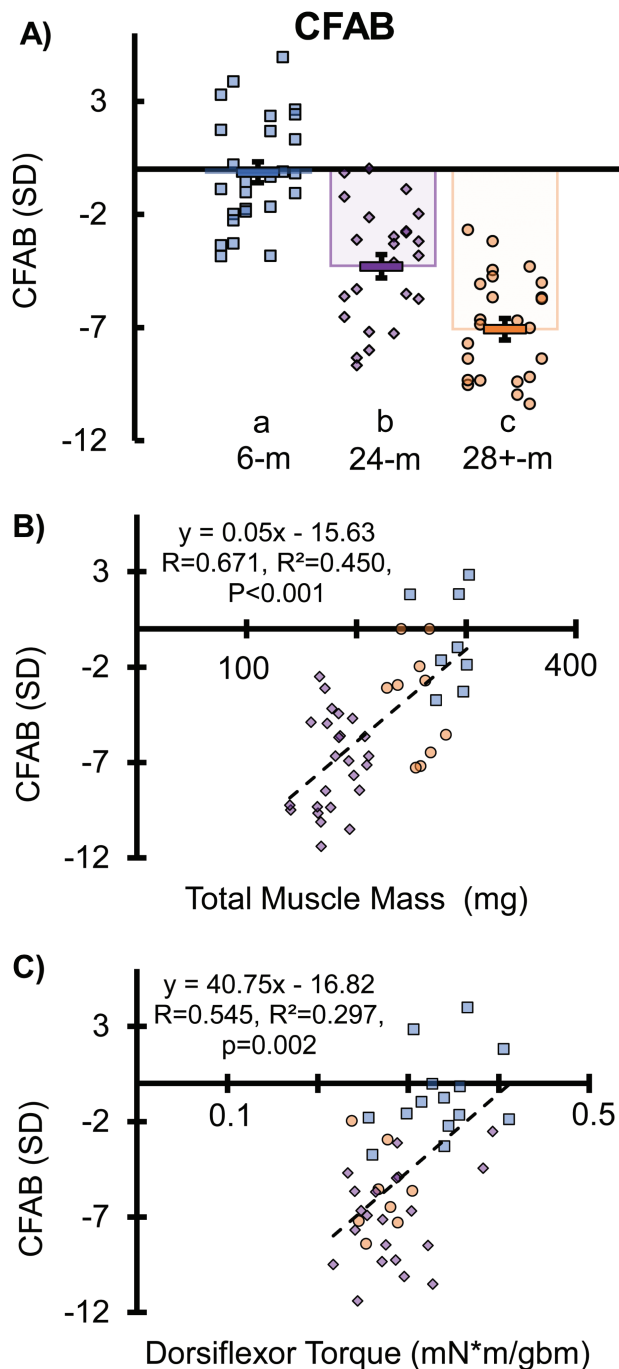

Figure 2. (**A**) Comprehensive Functional Assessment Battery (CFAB), (**B**) regression of sarcopenia outcome, total muscle mass versus CFAB, and (**C**) dorsiflexor torque versus CFAB. Key: (A) Different letters indicate a significant difference ($p < .05$) between groups (from analysis of covariance adjusted for body mass and using least significant difference post hoc test). Each symbol (squares for mice 6 months old, diamonds for 28+ months old, and circles for 28+ months old) indicates the test score for an individual mouse. The rectangles with error symbols (±SEM) in each grouping of ages indicate the mean value for that group. (B and C) Total mass (mg) is the sum of the mass of the plantaris, soleus, extensor digitorum longus, tibialis anterior, and gastrocnemius muscles. gbm = grams body mass; km = kilometers; mg = milligrams; mN = millinewtons; SD = standard deviation; SEM = standard error of measurement. Each symbol equals the regression of the scores of an individual mouse. Dashed line and equation equal the simple linear regression of the combination of all 3 groups. Full color version is available within the online issue.

the average 6 months and Supplementary Figure 1B illustrates that the 28+ months upper quartile is indistinguishable from the mean of 24 months. Thus, the highest functioning subpopulation in both older groups are equivalent in function to the average younger mouse.

Other Tests Not Part of CFAB

Body and muscle mass

The mean body mass (Supplementary Figure 2A) of the 6-month-old (30.96 ± 0.55 g) and 28+-month-old (31.28 ± 0.63 g) mice was not different ($p > .100$), but the 24-month-old mice (33.65 ± 0.10 g) were 9% heavier than the 6-month-old mice ($p < .001$) and 7% heavier than the 28+-month-old mice ($p = .008$). There was overall significant muscle mass reduction with age (Supplementary Figure 2B–F). Compared to the 6-month-old mice (171.90 ± 5.48 mg), the hearts of older mice (24-month old 209.10 ± 6.46 mg, +21.6%; 28+-month old 193.70 ± 6.81 mg, +12.7%) were heavier ($p < .050$, data not shown).

Contractile Physiology (Muscle Performance)

Ex vivo contractile physiology

The 24-month EDL (–32.6%, $p = .003$) and SOL (–37.2%, $p < .001$) maximum isometric force (P_0) was lower compared to the 6 months. The 28+-month EDL (25%, $p = .005$) and SOL (–37%, $p < .001$) P_0 was also lower compared to the 6 months. Soleus P_0 (–16.5%, $p = .020$) was lower in the 28+-month group compared to the 24-month group, though EDL did not change between 24 and 28+ months. Soleus PCSA was lower in both 24 (–15%, $p = .040$) and 28+-months (–14,%, $p = .013$) compared to the 6-months, but EDL PCSA was lower only in 28+-months (–23%, $p < .001$), although there was a tendency for the 24-month EDL to be smaller than the 6-months (–13%, $p = .060$) and the 28+-months to be smaller than the old (–11%, $p = .099$). Maximum specific force ($P_0/PCSA$) in EDL (–26%, $p = .036$) and SOL (–10%, $p = .025$ with outlier removed) was lower in 24-months versus 6-months, indicating that the changes in P_0 were due to both atrophy and muscle quality degradation. Interestingly, only the SOL-specific force was statistically lower in the 28+-months (–26%, $p < .001$) compared to the 6-months in this cohort subgroup (Supplementary Figure 3A–F). Supplementary Table 2 outlines the subgroup data for in vitro contractile physiology.

In vivo contractile physiology (dorsiflexor torque)

The 6-month-old mice (mean dorsiflexor torque = 0.333 ± 0.011 mN × m/gbm [grams body mass]) were stronger than both the 24-months (0.269 ± 0.003 mN × m/gbm, –19%, $p < .001$) and 28+-months (0.284±0.009 mN × m/gbm, –15%, $p < .001$). The 28+-month-old mice were 15% lower than 6-month-old mice ($p < .001$) but similar to 24-month-old mice ($p > .10$; Supplementary Figure 4).

Correlations and Regressions

We examined the relations between the variables with regression analysis (see Supplementary Material for details). Using the combined set of all mice of the 3 ages, we examined the relationship between the determinants of CFAB using linear regression (Supplementary Figures 5–7) and determined that there was no collinearity in the 5 determinants of the CFAB score. Supplementary Table 3 also illustrates the relationships of the determinants by age group. We also examined the relationship between CFAB and 2 determinants of sarcopenia (muscle mass and strength): total muscle mass ($R = 0.671$) and dorsiflexor torque ($R = 0.545$; Figure 2B and C). Dorsiflexor torque output was moderately correlated with total muscle mass ($R = 0.600$; Supplementary Figure 8). We determined how much physical function loss is associated with sarcopenia: notably 43.5% of variance. Furthermore, we determined that CFAB is predictive of age group based on logistical regression.

Power Analyses and Effect Sizes

In Supplementary Material, we detail the effect sizes (ES) of our CFAB determinants (VWR, rotarod, grip test, treadmill, and inverted cling) and power analyses of the data from the 6-month-old and 28+-month-old groups. In brief, we determined that overall in a power analysis of an independent *t*-test, CFAB produces a higher effect size with smaller number of animals with increased power compared to the individual determinants alone (VWR [$n = 8$, ES 1.52, actual power = 0.81], rotarod [$n = 5$, ES 2.05, actual power = 0.81], grip test [$n = 8$, ES 1.55, actual power = 0.82], treadmill [$n = 36$, ES 0.68, actual power = 0.81], inverted cling [$n = 18$, ES 0.97, actual power = 0.80], and CFAB [$n = 5$, ES 2.28, actual power = 0.88]). This indicates that CFAB is a successful composite scoring system that improves detectability of changes.

Discussion

The primary purpose of this study was to create a composite scoring system, CFAB, that would encompass the overall physical function of a mouse at a given point in time. Therefore, by basing this test on repeatable and noninvasive measurements, researchers can use CFAB in longitudinal or intervention studies, to investigate for example how exercise or pharmaceuticals affect age-related neuromuscular decline. Other potential uses include assessing the functional phenotype of transgenic mice to determine the physiological relevance of a gene, protein, or signaling pathway.

As expected, physical function declined overall with age in our mice. Previous work by our group and many others has also shown age-related muscle and functional decline (4,5,7,8,18–20). In prior studies, we created composite scoring systems to measure neuromuscular decline (the Neuromuscular Healthspan Scoring System [NMHSS]) and to detect changes in frailty status (the Frailty Intervention Assessment Value [FIAV]) (5,7,21). The NMHSS used multiple linear regression modeling to account for variability in the component measurements and vastly increased power of detection while simultaneously reducing the number of mice needed to reach 80% power, but utilized EDL in vitro contractile physiology (a complex terminal invasive procedure) limiting application and adoption (7). CFAB is loosely based upon the FIAV, which was designed to detect changes in frailty status, based on the reverse translation of the Fried Frailty Phenotype (5,6,18). The FIAV essentially uses a system similar to what we have described in CFAB (albeit with fewer unique tests) with one test before the intervention and a second test after, the difference between the two being the change attributed to the intervention. We believe that CFAB could be used longitudinally or as a within-group repeated measurement in the same manner and will test that hypothesis in future work, because, of course, the current study is cross-sectional. In Supplementary Material, we discuss in detail the relationship between CFAB, frailty, FIAV, and the NMHSS (Supplementary Tables 4 and 5). The CFAB scoring system is an evolutionary successor to these previous systems that we believe will serve as a valuable tool for future physiological investigation.

Functional Testing, CFAB, Sarcopenia, and Muscle Weakness

Sarcopenia leads to loss of physical function, muscle performance, and weakness (6,8,18,22). By examining the relationship between CFAB (physical function status) and muscle size/strength (total muscle mass and dorsiflexor torque), 43.5% of the variance was explained (adding torque in a multiple linear regression did not increase the linear relationship or predictability). As expected, there was age-related loss of EDL and SOL muscle mass and strength, further confirming evidence of sarcopenia in our model. Thus, the remaining variance beyond that associated with loss of muscle mass is likely caused by other factors—potentially neural, cardiovascular, or metabolic. For example, the older mice demonstrated significant evidence of cardiac hypertrophy, which we have shown in previous work to be related to lower functional performance in the rotarod (7). In future work, we will examine these other potential sources of functional loss in greater detail.

Relationships Among the Data

There was little evidence of collinearity between the measurements of the CFAB, indicating that the different tests measure different aspects of functional aging. Earlier work has also shown that the individual tests were not highly correlated with one another (5,7). Note that the logistic regressions were able to predict chronological age groups with a high degree of discernment, even though the upper quartile of the 24-month-old mice was statistically similar to the average 6-month-old mice and the upper quartile of 28+-month-old mice was statistically similar to the average 24-month-old mice, demonstrating stratification in the functional loss within age groups. There was also a positive association between overall muscle mass and CFAB. Thus, while sarcopenia is a major contributor to physical function decline and disability onset, there are other factors to consider at the physiological level that combine to alter functional aging trajectories within age groups which potentially help explain differential aging of the neuromuscular system.

Biological Versus Chronological Age: Functional Age

An emerging concept is that at the individual level, animals age at different rates overall and in different tissues (23–25). Functional status encompasses a juxtaposition of numerous physiological systems: neural, muscular, connective tissue, cardiovascular, pulmonary, vestibular, and metabolism, all of which alter with age (24,26–28). Individuals of the same age may have vastly different physical function and exercise capacity/tolerance profiles, potentially indicating an overall difference in fundamental aging rates within the age group. In other words, some individuals are "older biologically" than their chronological age would indicate (defined perhaps as >1 *SD* below the mean on CFAB), some are younger (>1 *SD* above the mean on CFAB), whereas others are average (middle range of mean ± 1 *SD*). Using CFAB to discern biological age versus chronological age, ie, functional age, in this fashion may allow the downstream mechanistic analysis of factors involved in the aging of various tissues by comparing genomic, epigenomic, transcriptomic, and proteomic data within and between age groups and levels of functional ability.

Caveats

One caveat is that our mice run less on running wheels than some other groups (29,30). We have observed lower numbers using these same running wheels in other studies and with other strains (5,31). We speculate that the difference may be the wheels, which are designed to be fully contained within the cages (for specific pathogen-free environments), and thus have a smaller diameter than some other wheel designs. This may make the wheels less comfortable to use for larger mice, encouraging less extensive running, which has been observed by others (30). However, as we are using these wheels to measure volitional exercise/activity rate, the principle behind their usage is valid because we are concerned with differences between groups not with actual values. Second, there were some violations of normality in the data (inverted cling and 28+-month wheel running). The violations in inverted cling were extensive enough to justify transforming (log10) the data before standardizing into CFAB scores. Both caveats warrant caution when attempting to generalize actual number test values between different cohorts.

In addition, we should note that because absolute outcome measure numbers may be influenced by the skill of the researcher to encourage the mice to perform, each test, if not all function, within a given study should optimally be performed by one individual. Furthermore, the selection of reference groups for other types of study designs using the CFAB concept should be carefully considered. While aging research often has an adult control (as the current study) for a reference group, the control group could just as easily be a non-exercise group in a training study, or CFAB could potentially be used within a group as a repeated measure with the first time point representing a comparison to the control baseline, and then changes documented by reapplying CFAB at distant time points (or pre- and post-intervention, for example) compared to that baseline score, as was done with FIAV (5).

Conclusions

As a composite scoring system, CFAB is highly sensitive and powerful, being able to detect differences in physical function between groups with increased power using a smaller sample size than the individual determinants alone, while also encompassing the discerning ability to detect potential frailty and neuromuscular decline with a good level of accuracy. Our comprehensive functional assessment battery as a numerical construct to quantify overall physical function status and exercise capacity will serve as the basis for future research into the molecular mechanisms underlying age-associated loss of physical ability.

The concept of resilience is an emerging important aspect of aging. Resilience is the ability of an organism to rebound back to normal levels after homeostatic perturbations. Comprehensive functional assessment battery may be used in future investigations examining how functional status is restored following a systemic challenge. Additionally, therapies designed to improve physical aptitude we predict could be measured for efficacy by comparing pre- and post-treatment CFAB scores.

One future step is to investigate the predictive capacity of CFAB to determine upcoming functional liability and thus be able to test interventions that may interdict the aging process by retarding the slope of the functional decay curve in advance of a future potential disability. To conclude, CFAB is a composite mathematical construct that provides a sensitive assessment of the ability of an individual mouse to perform complex motor tasks and exercise capacity in relation to the average of a control group (in our study, the 6-month-old mouse), which can provide a reliable, repeatable, noninvasive, and powerful numerical quantification to understand physical function.

Supplementary Material

Funding

Conflict of Interest

None declared.

Author Contributions

Conceptualization: T.G.G.; methodology: T.G.G.; validation: T.G.G. and R.M.; formal analysis: T.G.G. and R.M.; investigation: T.G.G., R.M., C.S.F., and C.R.B.; resources: T.G.G., C.S.F., and B.B.R.; writing—original draft: T.G.G.; writing—review and editing: T.G.G., B.B.R., C.S.F., C.R.B., and R.M.; supervision: T.G.G.; project administration: T.G.G.; funding acquisition: T.G.G. and C.S.F.

References

1. Rizzoli R, Reginster JY, Arnal JF, et al. Quality of life in sarcopenia and frailty. *Calcif Tissue Int*. 2013;93:101–120. doi:10.1007/s00223-013-9758-y

2. Marzetti E, Calvani R, Tosato M, et al. Sarcopenia: an overview. *Aging Clin Exp Res*. 2017;29(1):11–17. doi:10.1007/s40520-016-0704-5

3. Tieland M, Trouwborst I, Clark BC. Skeletal muscle performance and ageing. *J Cachexia Sarcopenia Muscle*. 2018;9:3–19. doi:10.1002/jcsm.12238

4. Graber TG, Fandrey K, Thompson LV. Novel individualized power training methodology: preserves physical function in adult and older mice. *Geroscience*. 2019;41:165–183. doi:10.1007/s11357-019-00069-z.

5. Graber TG, Ferguson-Stegall L, Liu H, Thompson LV. Voluntary aerobic exercise reverses frailty in old mice. *J Gerontol A Biol Sci Med Sci*. 2015;70:1045–1058. doi:10.1093/gerona/glu163

6. Seldeen KL, Lasky G, Leiker MM, Pang M, Personius KE, Troen BR. High intensity interval training improves physical performance and frailty in aged mice. *J Gerontol A Biol Sci Med Sci*. 2018;73:429–437. doi:10.1093/gerona/glx120

7. Graber TG, Ferguson-Stegall L, Kim JH, Thompson LV. C57BL/6 Neuromuscular Healthspan Scoring System. *J Gerontol A Biol Sci Med Sci*. 2013;68:1326–1336. doi:10.1093/gerona/glt032

8. Graber TG, Kim J-H, McLoon LK, Grange RW, Thompson LV. C57BL/6 lifespan study: age-related declines in muscle power production and contractile velocity. *Age*. 2015;37:36. doi:10.1007/s11357-015-9773-1

9. Neelakantan H, Brightwell CR, Graber TG, et al. Small molecule nicotinamide N-methyltransferase inhibitors activate senescent muscle stem cells and improve regenerative capacity of aged skeletal muscle. *Biochem Pharmacol*. 2019;163:481–492. doi:10.1016/j.bcp.2019.02.008

10. Graber TG, Rawls BL, Tian B, et al. Repetitive TLR-3-mediated lung damage induces skeletal muscle adaptations and cachexia. *Exp Gerontol*. 2018;106:88–100. doi:10.1016/j.exger.2018.02.002

11. Bradley JV. Robustness? *Br J Math Stat Psychol*. 1978;31:144–152. doi:10.1111/j.2044-8317.1978.tb00581.x

12. Ghasemi A, Zahediasl S. Normality tests for statistical analysis: a guide for non-statisticians. *Int J Endocrinol Metab*. 2012;10(2):486–489. doi:10.5812/ijem.3505

13. Blanca MJ, Alarcón R, Arnau J, Bono R, Bendayan R. Non-normal data: is ANOVA still a valid option? *Psicothema*. 2017;29:552–557. doi:10.7334/psicothema2016.383

14. Sullivan GM, Feinn R. Using effect size-or why the P value is not enough. *J Grad Med Educ*. 2012;4(3):279–282. doi:10.4300/JGME-D-12-00156.1

15. Coe R. It's the effect size, stupid: what "effect size" is and why it is important. Presented at the 2002 Annual Conference of the British Educational Research Association, University of Exeter, Exeter, Devon, England; September 12–14, 2002. http://www.leeds.ac.uk/educol/documents/00002182.htm. Accessed April 10, 2019.

16. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*. 2007;39:175–191. doi:10.3758/bf03193146

17. Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav Res Methods*. 2009;41:1149–1160. doi:10.3758/BRM.41.4.1149

18. Thompson LV. Age-related muscle dysfunction. *Exp Gerontol*. 2009;44:106–111. doi:10.1016/j.exger.2008.05.003

19. Justice JN, Cesari M, Seals DR, Shively CA, Carter CS. Comparative approaches to understanding the relation between aging and physical function. *J Gerontol A Biol Sci Med Sci*. 2016;71:1243–1253. doi:10.1093/gerona/glv035

20. Justice JN, Carter CS, Beck HJ, et al. Battery of behavioral tests in mice that models age-associated changes in human motor function. *Age (Dordr)*. 2014;36:583–592. doi:10.1007/s11357-013-9589-9

21. Liu H, Graber TG, Ferguson-Stegall L, Thompson LV. Clinically relevant frailty index for mice. *J Gerontol A Biol Sci Med Sci*. 2014;69:1485–1491. doi:10.1093/gerona/glt188

22. Janssen I, Heymsfield SB, Ross R. Low relative skeletal muscle mass (sarcopenia) in older persons is associated with functional impairment and physical disability. *J Am Geriatr Soc*. 2002;50:889–896. doi:10.1046/j.1532-5415.2002.50216.x

23. Melis JP, Jonker MJ, Vijg J, Hoeijmakers JH, Breit TM, van Steeg H. Aging on a different scale – chronological versus pathology-related aging. *Aging (Albany NY)*. 2013;5:782–788. doi:10.18632/aging.100606

24. Khan SS, Singer BD, Vaughan DE. Molecular and physiological manifestations and measurement of aging in humans. *Aging Cell*. 2017;16:624–633. doi:10.1111/acel.12601

25. Ferrucci L, Gonzalez-Freire M, Fabbri E, et al. Measuring biological aging in humans: a quest. *Aging Cell*. 2020;19:e13080. doi:10.1111/acel.13080

26. Li Q, Wang S, Milot E, et al. Homeostatic dysregulation proceeds in parallel in multiple physiological systems. *Aging Cell*. 2015;14(6):1103–1112. doi:10.1111/acel.12402

27. Cohen AA. Complex systems dynamics in aging: new evidence, continuing questions. *Biogerontology*. 2016;17:205–220. doi:10.1007/s10522-015-9584-x

28. Picca A, Calvani R, Bossola M, et al. Update on mitochondria and muscle aging: all wrong roads lead to sarcopenia. *Biol Chem*. 2018;399:421–436. doi:10.1515/hsz-2017-0331

29. Lee JY, Paik IY, Kim JY. Voluntary exercise reverses immune aging induced by oxidative stress in aging mice. *Exp Gerontol*. 2019;115:148–154. doi:10.1016/j.exger.2018.08.009

30. Manzanares G, Brito-da-Silva G, Gandra PG. Voluntary wheel running: patterns and physiological effects in mice. *Braz J Med Biol Res*. 2018;52:e7830. doi:10.1590/1414-431X20187830

31. Graber TG, Fry CS, Brightwell CR, et al. Skeletal muscle-specific knockout of DEP domain containing 5 protein increases mTORC1 signaling, muscle cell hypertrophy, and mitochondrial respiration. *J Biol Chem*. 2019;294:4091–4102. doi:10.1074/jbc.RA118.005970