
Article:
Bellantuono, I. orcid.org/0000-0001-9994-6987, de Cabo, R., Ehninger, D. et al. (9 more authors) (2020) A toolbox for the longitudinal assessment of healthspan in aging mice. Nature Protocols, 15 (2). pp. 540-574. ISSN 1754-2189

https://doi.org/10.1038/s41596-019-0256-1

A toolbox for the longitudinal assessment of healthspan in ageing mice

Contributing authors

I Bellantuono[1*], R de Cabo[2], D Ehninger[3], C Di Germanio[2], , A Lawrie[4], J Miller[5], SJ Mitchell[6], I Navas-Enamorado[2], PK Potter[7], T Tchkonia[5] , JL Trejo[8], DW Lamming[9,10*]

[1]MRC–Arthritis Research UK Centre for Integrated research into Musculoskeletal Ageing, University of Sheffield, Sheffield, UK

[2]Translational Gerontology Branch, National Institutes of Health, Baltimore, MD, USA. - clara.digermanio@nih.gov, nach.navas-enamorado@nih.gov; decabora@grc.nia.nih.gov;

[3]German Centre for Neurodegenerative Disease, Bonn, Germany - Dan.Ehninger@dzne.de

[4]Department of Infection, Immunity & Cardiovascular Disease, University of Sheffield, Sheffield, UK - a.lawrie@sheffield.ac.uk

[5]Robert and Arlene Kogod Center on Aging, Mayo Clinic, Rochester, MN, USA. - Tchkonia.Tamar@mayo.edu, Miller.Jordan@mayo.edu

[6]Harvard T.H. Chan School of Public Health, Boston, Massachusetts 02115, USA. smitchell@hsph.harvard.edu

[7]Oxford Brookes University, Oxfordshire, UK. ppotter@brookes.ac.uk

[8]Cajal Institute (CSIC), Madrid, Spain - jltrejo@cajal.csic.es

[9]William S. Middleton Memorial Veterans Hospital, Madison, WI, USA

[10]Department of Medicine, University of Wisconsin-Madison, Madison, WI, USA - dlamming@medicine.wisc.edu

*****Corresponding authors

Ilaria Bellantuono,
Professor in Musculoskeletal Ageing
MRC/Arthritis Research-UK Centre for Integrated Research into Musculoskeletal Ageing
Department of Oncology and Metabolism
D Floor
The Medical School
Beech Hill Road
Sheffield S10 2RX
Telephone: +44 (0)114 2159037
Fax: +44 (0)114 271 2475

Dudley Lamming
Assistant Professor of Medicine
University of Wisconsin-Madison
2500 Overlook Terrace
VAH C3127 Research 151
Madison, WI 53705, USA
Telephone: +1 608-256-1901 x12861
Fax: +1 608-263-9983

Keywords: aging, healthspan, phenotyping, geroprotectors, cardiovascular function, cognitive function, neuromuscular function, metabolic health, glucose tolerance test, insulin tolerance test, body composition analysis

EDITORIAL SUMMARY A series of techniques (echocardiography, novel object recognition, grip strength, rotarod, glucose and insulin tolerance tests, body composition, and energy expenditure) are used to assess the health of mice.

TWEET Assessing the health of aging mice

COVER TEASER Aging mice health assessment

RELATED LINKS

Key reference(s) using this protocol

Martin-Montalvo, A. *et al.* Metformin improves healthspan and lifespan in mice. *Nature communications* **4**, 2192, doi:10.1038/ncomms3192 (2013)
Neff, F. *et al.* Rapamycin extends murine lifespan but has limited effects on aging. *J Clin Invest* **123**, 3272-3291, doi:10.1172/JCI67674 (2013).
Zhu, Y. *et al.* The Achilles' heel of senescent cells: from transcriptome to senolytic drugs. *Aging Cell* **14**, 644-658, doi:10.1111/acel.12344 (2015).

Abstract

The number of people aged over 65 is expected to double in the next 30 years. For many, living longer will mean spending more years with the burdens of chronic diseases such as Alzheimer's, cardiovascular disease, and diabetes. Although researchers have made rapid progress in developing geroprotective interventions that target mechanisms of ageing and delay or prevent the onset of multiple concurrent age-related diseases, a lack of standardized techniques to assess healthspan in preclinical murine studies has resulted in reduced reproducibility and slowed progress. To overcome this, major centres in Europe and the USA skilled in healthspan analysis came together to agree upon a toolbox of techniques which can be used to consistently assess the healthspan of mice. Here, we describe the agreed toolbox which contains protocols for echocardiography, novel object recognition, grip strength, rotarod, glucose and insulin tolerance tests, body composition, and energy expenditure. They can be performed longitudinally in the same mouse over a period of 4-6 weeks to test how candidate geroprotectors affect cardiac, cognitive, neuromuscular and metabolic health.

Introduction

As the average age of the world population increases, for many living longer will not mean living healthier. The number of healthy years is predicted to increase at a slower pace or not at all compared to lifespan. The result is that people are expected to spend an increasing number of years suffering from the burdens of multiple chronic age-associated diseases such as heart disease, Alzheimer's disease, and diabetes. Over 60% of those aged 65 and over have more than one chronic condition (multimorbidity), often leading to frailty (an accumulation of multiple deficits and loss of resilience to adverse events) and a resulting loss of independence[1]. This continuously growing trend will mean a substantial increase in health and social care costs[1]. For example the European Commission has estimated that every year of health lost will cost to the European Union 0.3 percentage point of GDP, equivalent to 54 billion euros[1].

For this reason, studies on how to maintain health with age and how to compress the period of multimorbidity have become an emerging priority[2]. Recent research suggests that ageing, the major risk factor for the development of many deleterious chronic conditions, may be modifiable by both lifestyle (e.g., diet, exercise) and pharmacological interventions (geroprotectors), which in turn may be able to simultaneously delay the onset of multiple age-related diseases and reduce the risk of developing multimorbidity and frailty[1]. This has stimulated great interest in developing interventions to promote healthspan[3-6].

Mice are the model organisms most often used for preclinical analysis of interventions due to their mammalian physiology, which resembles that of humans in many aspects, the relatively inexpensive cost to house and treat the animals, and their relatively short lifespan. However, measuring healthspan (the portion of life that is relatively healthy and free from major deficits that impair quality of life) is challenging because of the multiple parameters which need to be assessed. Despite this difficulty, it is widely agreed that measuring healthspan comprehensively is imperative[1,7]. Maximizing the robustness, reproducibility, and utility of these measurements of healthspan requires a consensus on which outcomes should be measured, the ages at which mice should be examined, and the frequency of examination.

At present the majority of laboratories do not comprehensively assess healthspan, and testing is not all-inclusive in relation to organ systems but rather focused on one or two specific organs of choice. Further, there are differences in the experimental design and the way the techniques are performed. In combination, these result in both a knowledge gap about the effectiveness of candidate geroprotective agents, and a lack of reproducibility. Therefore, a consensus on the best practices to assess healthspan and development of a standard operating procedure (SOP) is required.

Development of the toolbox

In an attempt to overcome these limitations, experts in healthspan assessment of mice gathered to compare practices in their laboratories. The toolbox was developed via a series of meetings and conference phone calls, culminating in in-person meeting at a MouseAGE/INFRARONTIER-sponsored workshop in Athens, Greece in October of 2017. Here, we present the consensus toolbox for the measurement of healthspan in ageing mice using a battery of tests, which assess the major organ systems in the same animals longitudinally. The toolbox consists of protocols assessing cardiac function (echocardiography), muscle strength and neuromuscular function (grip strength, hanging bar, and rotarod), metabolic health (glucose and insulin tolerance test, body composition, and energy balance), and cognitive function (novel object recognition). We also recommend mice are assessed for frailty. The protocols in this toolbox can be repeated longitudinally during the lifespan of a mouse, or during the course of treatment of aged mice to which geroprotective agents have been administered. At the time of euthanasia, collection of

tissues for histological analysis is also recommended for studies with a planned euthanasia age.

The assays were chosen based on the evaluation of a number of criteria: clinical relevance, good reproducibility, a low level of invasiveness or exposure to radiation and stress for the animal (e.g. no need for anaesthesia when possible), and a low level of technical difficulty.

For the evaluation of cardiac function, echocardiography was selected over blood pressure measurement despite the need for anaesthesia and for specialized personnel to perform the analysis because the latter procedure has very high variability, likely due to stress, and there is significant difficulty in reproducing results.

In the case of neuromuscular function, we chose a panel of easy to perform tests with very low impact on animal welfare. Among these assays, grip strength, cage top and hanging wire were chosen over the testing of forced treadmill running capacity (FTR). Wire hang is also known as wire suspension latency test, wire hang, hang test or body suspension in the literature. This combination of tests was chosen because they are less demanding for the aged animals than the treadmill running test. Aged animals may be lost during the latter and therefore it is preferable to perform it on a dedicated cohort of animals. We have also chosen rotarod because, in addition to serving as another measure of endurance, it also assesses coordination and balance. Alternative assays such as beam balance and beam walking assay are less effective at detecting mild impairments than rotarod [8]. When performing grip strength it is worth remembering that it has the significant limitation of being very sensitive to the ability of the operator. In addition, several of these assays are heavily influenced by the weight of the animal. For this reason, the evaluation of the neuromuscular system relies on more than one test, and where possible, it is advisable to verify conclusions by post-mortem histological analysis of muscle. We also suggest to perform follow-up studies if there is any evidence of a change in muscle function, using FTR and evoked leg muscle strength and fatigue [3,5,9,10].

For metabolic function, glucose and insulin tolerance tests were selected to measure glucose homeostasis over the hyperinsulinemic-euglycemic clamp due to the requirement for surgery and euthanasia following the clamp procedure. This precludes longitudinal assessment and is very stressful to the animals. We selected intraperitoneal (IP) dosing of glucose due to the ease of training individuals and the consideration that oral glucose delivery is considered more stressful. Body composition analysis by MRI was selected over DXA due to the requirement for anaesthesia, the speed of measurement, and accuracy, as MRI provides a better estimate of both fat mass and lean tissue mass [11]. Metabolic chambers to assess energy balance via indirect calorimetry, activity tracking, and food intake measurement are the gold standard method to track energy balance and can be used longitudinally with a minimum of stress to the animals. They were selected over alternative apparatus that can only track a subset of these measurements, and thus provide only a partial snapshot of energy balance.

A separate session was dedicated to the choice of behavioural tests. Consensus versions of protocols for each technique were modified from those already published by the

individual authors[3-6,12-21]. A key requirement for cognitive assays suitable for longitudinal analyses is that order effects need to be avoided as much as possible. A number of commonly used cognitive tasks in rodents, such as the Morris water maze or contextual fear conditioning, have limited value for longitudinal assessments because performance on these assays will be strongly influenced by previous experiences in these tasks. We chose novel object recognition as it is well suited for multiple longitudinal assessments of the same sets of mice using different object pairs. Other behavioural assays, such as the recently developed touchscreen-based automated battery system [22], would also be suitable for longitudinal assessments of cognitive function. However, novel object recognition has particularly low demands regarding the availability of specialized equipment and can hence be readily implemented in a broad range of laboratories. Moreover, novel object recognition places minimal stress on the animals tested as it does not involve exposure to aversive stimuli, food restriction, single housing or extensive periods of pre-training. More detailed explanations of novel object recognition can be found in Benice and Raber (2008)[23].

We also recommend uses assess mice according to the frailty index, based on the Rockwood frailty score including 31 different parameters, which we do not describe here due to the extremely detailed description available from the developers [24-26]. This is our preferred method to score frailty compared to measuring frailty using a Fried Frailty Criteria-like score because the latter is based on the assessment of only 4 criteria (grip-strength, walking speed, physical activity, and endurance) and was considered more a measure of musculoskeletal fitness rather than an accumulation of deficits in several organ systems. The Rockwood-like frailty index better captures the potential benefits of geroprotectors across multiple systems.

At the end of the experiments we recommend harvesting as many tissues as possible, as well as blood for clinical biochemistry and to perform analysis of bone, e.g. by X-Ray microCT. It was decided that analysis of bone would be done only after terminal anaesthesia to avoid potential detrimental effects due to radiation exposure. Longitudinal analysis of bone requires exposure of mice to repetitive doses of radiation and anaesthesia. Although the dose received at each imaging session is small and only in the region of interest, the cumulative effects are unknown.

An alternative approach to the testing of healthspan was proposed by Sukoff-Rizzo *et al*. [27] However, this approach was primarily focused on the in-depth study of motor, sensory, and cognitive function, and does not evaluate important parameters of healthspan such as cardiac and metabolic function [27]. Considering the longitudinal nature of these studies and considering that some of the tests reported by the protocol of Sukoff-Rizzo are taxing for ageing animals, we advise investigators to perform some of the tests described in this protocol in a second, parallel cohort of animals when alterations in neuromuscular function or behaviour are found using our pipeline.

Using the procedures included in this toolbox, our laboratories have successfully characterized the effectiveness of several compounds, including metformin[18], nicotinamide[20], rapamycin[4,6], and senolytics[28,29], as well as other interventions including diet[30-32] and exercise[33] as geroprotective agents on mice, and determined the effect of specific genetic modifications [12,19,34] on the health and longevity of mice. Consensus across laboratories was reached by performing a comparison of the protocols in use in each laboratory, together with evidence of best practice and consideration of animal welfare, particularly when animals reach old age.

Application of the toolbox

Healthspan assessment can be used when studying gene function using genetically modified animals, or for testing new compounds (e.g. geroprotectors), diet or exercise regimens. For example, we have used this toolbox to identify new genes involved in the development of age-related diseases, which can be used as novel drug targets[12]. In addition, we have studied the efficacy, side effects, and mechanism of action of geroprotectors using several tests in this toolbox. These drugs have emerged for their ability to delay the onset of multiple concurrent age-related diseases and boost resilience by modulating mechanisms of ageing such as senescence, autophagy, and inflammation[1]. More than 200 compounds have been described as geroprotectors (http://geroprotectors.org/). Our studies, collectively, have shown that drugs such as rapamycin, resveratrol, metformin, senolytics (e.g. fisetin, dasatininb, and querectin[35,36]) which remove senescent cells, or dietary interventions can slow the development of cataracts, osteoarthritis, osteoporosis, the loss of muscle mass and can improve cardiac function[3-6,13-21,35-42]. Similarly, we have shown that decreased consumption of protein or of specific dietary amino acids improves metabolic health[31,43]. We recommend the use of this toolbox for a first assessment of healthspan whenever testing geroprotectors or assessing the function of specific genes in healthy ageing. These studies can serve as a guide for a subsequent, in-depth assessment of specific organ systems and healthspan effects.

The toolbox has been designed with C57BL/6J mice in mind because of the amount of data already available for this commonly used strain. However, with some modification it can be used for other strains such as UM-HET3 mice (the F2 progeny of (BALB/cJ x C57BL/6J) mothers and (C3H/HeJ x DBA/2J) fathers; parental strains from The Jackson Laboratory). Some procedures have also been tested on mice of different genetic backgrounds, including PWK/PhJ (The Jackson Laboratory) and C57BL/6NHsd, DBA/2NHsd, 129P2/OlaHsd, and C3H/HeNHsd (Envigo), and many of the toolbox procedures have been tested on genetically modified mice, including dwarf and progeroid animals [15,34,44-46].

Comparison with other approaches

To date, the vast majority of healthspan studies have taken an *ad hoc* approach to the collection of longitudinal data that has been limited by the interests of the investigators in one or two specific tissues. This has limited the amount of data collected in a single study, slowed scientific advance overall and limited the translation of geroscience discoveries. For example, it was only recently shown that short-term calorie restriction and rapamycin can rejuvenate the aging mouse heart, reversing pre-existing age-dependent cardiac hypertrophy and diastolic dysfunction[47,48]. If these types of studies were routinely performed during longevity studies, the beneficial effects of rapamycin on cardiac diastolic function would have been identified several years earlier. For example, while there have been several studies demonstrating that late-life treatment with rapamycin extends lifespan[6,49-51], none of these studies examined cardiac function. Likewise, there are studies testing geroprotectors and focusing on cognition and metabolic health, showing beneficial effects on cognition in mouse models of Alzheimer's disease or aged mice[52,53], and improving the metabolic health of aged mice[20,29,54]. For the greater part, however, those studies that examined cognition did not assess metabolic health, and vice-versa. It would also be interesting to know if the beneficial effects of these short-term interventions persist or are temporary.

The major problem of older age is that "it comes as a package"[55] and the major health challenges are due to the accumulation of deficits and the presence of concurrent diseases. Therefore, a systematic and holistic approach which goes beyond the study of the single disease is required. In addition the use of a standardised toolbox allows the

comparison of measure across laboratories, improving reproducibility. For example, the effects of interventions including intermittent fasting and rapamycin have alternatively been reported to either have minimal or significant effects on aging and healthspan [4,32,56,57]; it is difficult to determine whether the differences are due to the differences in regimen or to the specific endpoints considered. While a few studies have provided a comprehensive assessment of healthspan measures across different physiological systems[4], these studies employed cross-sectional designs, and were not suited for an examination of the longitudinal development of health measures during ageing in individual mice. A principle advantage of the toolbox here is its simplicity, ready training of investigators, the minimal stress placed on the animals, and its applicability even to very old mice[29]. While there has been concern about phenotyping of mice possibly altering their lifespan, we have found that mice subject to longitudinal metabolic phenotyping, for example, have a similar lifespan to wild-type C57BL/6J mice not subject to such studies [6,30,58]. Overall, use of the toolbox will speed the prioritization of possible geroprotectors for further preclinical and clinical testing.

Experimental design

When testing geroprotectors, a minimum of two groups of mice is required: a group treated with the geroprotector and a vehicle-treated control group; Treating young mice is also often desirable to control for effects which are independent of age or may be detrimental such those seen with rapamycin[4,59]. When testing genetically modified animals, there may also be a need to include additional control groups based on how the mice were generated and the precise experimental question to be answered. Every intervention should be tested in male and female mice as there are many cases of sexual dimorphism[60-63]. In our experience, groups of at least 20 animals per sex are required to detect physiologically relevant changes in assays included in this toolbox, such as rotarod performance, glucose tolerance, and grip strength[4,6,16,17,31,43,64,65]. These have been obtained using in house and published data on mean and standard deviation for each test, using biological relevant effect sizes with a power of 80% and alpha of 0.05. We have considered the same size of effect following exercise, an intervention known to have positive effects in humans. We recommend that animals be randomly assigned to groups at the cage level, using computing generated assignment of mice to the treated and control group and stratified by weight; we typically observe that stratification by weight is sufficient to match most baseline measures of health in young wild-type animals. Stratification based also (or instead) on frailty may be considered when group assignment is performed in aged mice to reduce variability and group size.

We also recommend blinding of the experiments, i.e. that the allocation of the treatment groups is performed by a different researcher and should not be revealed to the people conducting the experiment until the data have been analysed. This is particularly important for tests such as frailty or grip strengths, which are operator dependent. It may not always be possible as some phenotypes (i.e. mice treated with calorie restriction) are visually obvious.

For both cross-sectional and longitudinal deaths, it is important to include sufficient numbers of mice to allow for age-related mortality, particularly after 18 months of age. In the hands of some of the investigators, female mice on a C57BL/6J genetic background or from the NIA Aging Mouse Colony experience 30-36% mortality between 20 and 27 months of age (50-66% mortality between 20 and 30 months of age), while male mice experience 8-18% mortality between 20 and 27 months of age (and 30-36% mortality between 20 and 30 months of age) ([6] and data not shown). The overall mortality rate also varies based not only on strain and genetic background, but also diet and animal facility; for example, the NIA's Intervention Program has found a persistent difference in lifespan

between mice test sites despite standardized mice, diet, and husbandry[66]. If local data for the survival of a strain is not available, the mortality between any two ages can be estimated for many strains using lifespan data published by the NIA and by The Jackson Laboratory[58,67].

Fig. 1 is a schematic representation of the experimental design for testing in C57BL/6J mice, the frequency of assessment and the type of tests. The sequence with which the tests are performed is also important and the time in between tests to avoid stress to the animals as much as possible. We recommend to perform the multiple assays over a period of 4-6 weeks using the following order, novel object recognition, frailty index, grip strength, cage top and hanging bar, followed by rotarod, body composition, energy expenditure, GTT, ITT and echocardiography.

Limitations of the approach Testing healthspan involves complex experiments, which are costly and requires an in-depth knowledge of multiple organ systems combined with knowledge of the ageing process and welfare issues in mice. The protocols described in this toolbox summarize the knowledge collected from years of experience in the area of healthspan assessment and murine ageing biology. Their availability will reduce animal usage and costs by reducing the need to optimize each technique when establishing healthspan assessment platforms in other laboratories. It will also ensure that the data are comparable and reproducible across laboratories, speeding up progress.

Regulatory requirements. These vary in each country and between institutions, and therefore it is always advisable to check with the relevant veterinary authorities, and gain regulatory approval as may be required, before performing any procedure. At each institution, all authors performing any of the described procedures received regulatory approval and oversight as appropriate. This included approval/oversight by the UK Home Office (IB, PP), the Institutional Animal Care and Use Committee of the William S. Middleton Memorial Veterans Hospital (DL), Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen, Recklinghausen, Germany (DE), Mayo Clinic Institutional Animal Care and Use Committee (TT, JDM), the Committee of Ethics and Animal Experimentation of the Cajal Institute, Ethics Committee of the CSIC, and the Animal Protection Area of the Ministry of Environment of the Community of Madrid (JLT), and the Animal Care and Use Committee of the National Institute on Aging, NIH, US (RdC, CDG, SJM, IN).

Requirements for staffing, expertise, and equipment. While the majority of the techniques in the toolbox do not require an abundance of highly specialized staff, training is required, particularly for the grip strength test and echocardiography. It is advisable that PhD students and post-docs visit a lab which is running these tests routinely. This is particularly important for grip strength as this test is highly operator dependent. For echocardiography a specialized facility and extensive specialized training of staff is required for the acquisition, analysis and interpretation of the images. Some techniques such as the glucose tolerance test require a significant investment of hands-on staff time. Likewise, performing several informative protocols in the toolbox, including determination of body composition, cardiac function, and energy balance, requires an upfront investment in single-purpose equipment, and an ongoing commitment to maintain that equipment.

Handling of aged mice. Working with aged animals presents additional difficulties when compared to working exclusively with young animals. Animal care, laboratory, and veterinary staff should be familiar with the care and handling of aged mice, and be able to distinguish between an old mouse, a sick mouse and an old, sick mouse. Older mice tend to be less active, may lose weight, and appear less healthy, and humane endpoints developed for

younger mice may not be appropriate for older mice. With the time, expense and ethical implications of ageing cohorts of mice it is important that aged mice are not culled unnecessarily, thus affecting cohort size, or allowed to exceed humane endpoints or ethical standards. Staff should become familiar with how the strain(s) or line(s) age, their care, and the treatment of spontaneous phenotypes. The age of breeding stock can have effects on the offspring[13,68-72]; therefore knowledge of the age of breeding mice and even limited phenotyping such as body composition or weight, would be useful metadata. Ideally, the age of breeding stock would be standardised to minimise variation in offspring. Phenotypic data in aged mice tends to exhibit a greater variance than that in young mice[65], which may reflect the differences in ageing between individuals, even within inbred strains. With the increased variance of data from aged mice the standardisation of protocols down to the finest detail is critical as is the recording of metadata as small variations can be amplified thus confounding the output of experiments. This includes time of day of phenotyping and the operator.

Anaesthesia. The cardiac function protocol requires anaesthesia. Repeated anaesthesia – as would be required for longitudinal examination of cardiac function in aged mice - has been suggested to be linked to cognitive impairment in humans. However, repeated anaesthesia with isoflurane does not result in prolonged cognitive deficits at least in mice that are approximately 18 months old[73]. In this toolbox, the animal is euthanized after the last echocardiography at 24 months of age. Aged animals may also be more likely to have negative reactions to anaesthesia[74-76].

Timing may vary. Timings and frequency are for ageing C57BL/6J mice, but may vary for other strains or genetically modified models. For example, it may prove appropriate to begin phenotyping at an early age for very short-lived mice or for genetically modified mice expected to begin developing phenotypes at an early age[34,44]. Conversely, when mice are expected to develop phenotypes at an advanced age, or an intervention is begun only late in life, it is likely a more efficient approach to not begin phenotyping mice until later in their lifespan or immediately prior to the start of an intervention[6,47,49].

Suitability for progeroid mice. Many of the metabolic phenotyping protocols in the toolbox are not suitable, or are suitable only with modification, for the use of progeroid mice, which are often small, extremely frail, and prone to hypothermia if singly housed. When using metabolic chambers, in which animals are single-housed without bedding, some investigators suggest that animals which may be more sensitive to cold stress be analysed for a shorter time period than suggested below, typically approximately 24-26 hours, or the temperature in the room or specifically in the metabolic chamber (if housed in a temperature controlled cabinet) can be raised to prevent hypothermia.

Limitations of intraperitoneal administration of glucose. While we have provided instructions below for performing a glucose tolerance test using glucose administered via intraperitoneal injection (IPGTT), glucose can also be administered via oral gavage (OGTT). An OGTT is likely more physiological since glucose is administered orally, and incretin hormones are engaged. Incretin hormones (GIP and GLP1) are secreted from the intestine in the blood stream and influence the insulin secretory response and pancreatic beta cells; they also have effects on glucagon response and fat deposition[77-79]. If investigators are specifically asking questions that involve incretins, it is likely they will need to perform OGTTs.

However, drawbacks to OGTT include that it needs to be performed by a very well trained operator to reduce stress to the animals, which may influence the outcome of the study. Further, in some countries, oral gavage is considered a highly invasive procedure for the

animal and needs to be very well justified for its use. In contrast, IPGTT is a technique that can be rapidly taught and involves minimal stress for the animals. Finally, IPGTT and OGTT provide comparable results in young and aged C57BL/6J mice [80]. The majority of our labs administer IPGTTs unless a specific question requires the use of OGTT.

Limitations of the novel object recognition task. It involves only a single training session[81,82] and hence it is not possible to resolve potential effects on learning rate which would require repeated training. It should also be noted that the level of object exploration by the animals can be low or inconsistent which may require the exclusion of animals in order to avoid skewing of results. Also, it is necessary to carry out careful pilot studies to ensure that training period, delay intervals and test duration are optimized for suitable behavioural performance within a given experimental context. Moreover, one needs to carefully evaluate all object pairs used in the task to ensure that there is no intrinsic difference in object exploration by the animals. Conditions should always be counterbalanced to minimize any effect of differential intrinsic object preference.

The influence of the environment. The standardization of protocols is critical to minimize variation due to environmental differences or gene-environment interactions. Recent studies have highlighted how apparently minor differences, including timing, temperature, bedding, as well as variation in the microbiome can affect the outcome and reproducibility of mouse experiments[83]. The experience of the National Institute on Aging's Interventions Testing Program has demonstrated that even when mouse source, diet, and husbandry conditions are intentionally coordinated to be identical, there remain site-specific variations in murine lifespan[84]. While these studies demonstrate the difficulty of achieving perfect reproducibility in experiments conducted at disparate sites, they also demonstrate that by matching known variables as closely as possible, reproducibility can be enhanced by minimizing differences. Thus, we encourage investigators to control for those variables that are most likely to impact the experiment, such as the operator, the testing room, the time of the day, and the equipment utilized. Some factors such as housing conditions (number of animals per cage, type of enrichment, and noise) or diet may not always be under the control of individual researchers because it is determined at an institutional level. These factors should be carefully recorded as metadata along with experimental data.

MATERIALS

REAGENTS

Laboratory mice

The procedures described here were primarily developed with mice on a C57BL/6 genetic background (specific variants include but are not limited to C57BL/6J from The Jackson Laboratory, C57BL/6J from Janvier Labs, C57BL/6N from Taconic Biosciences, and C57BL/6J.Nia from the National Institute on Aging Aged Rodent Colony).

The procedures can be readily carried in healthy animals of a wide range of ages; cautions regarding risks of specific procedures to old animals (C57BL/6J mice over approximately 24 months of age) are noted in specific procedures.

CAUTION: All experiments should be performed in accordance with relevant guidelines and regulations and approval of the relevant institutional committees.

General reagents

- Cleaning supplies

CRITICAL All of the equipment utilized in the procedures below must be cleaned after use, using procedures suitable to the equipment and in accordance with both the manufacturer's instructions and institutional policies.

- Paper towels

Cardiac function

- Isoflurane for use in isoflurane vaporizer

CAUTION: Isoflurane is a respiratory depressant and chronic exposure can negatively impact health, resulting in effects including hypotension, tachycardia, respiratory depression, and elevated blood glucose level. Isoflurane should be used in combination with equipment designed to minimize exposure (e.g. vaporizer with an activated carbon scavenging system), and institutional standard operating procedures to ensure safe use should be followed.

- Medical Oxygen (e.g. BOC) or Air source for isoflurane vaporiser
- Conductive electrocardiography (ECG) Gel
- Aqueous medical ultrasound gel (e.g. AquaSonic)
- Micropore Tape (e.g. 3M)
- Electric small animal hair clippers
- Hair removal lotion (Veet)

Energy Balance

- Respiratory gasses. Gasses for the calibration of the O_2 and CO_2 sensors will be specified by the equipment vendor, but must be ordered from gas vendors.

Glucose tolerance test

30% glucose: Prepared by dissolving 30% (30g/100mL) glucose (G7021, MilliporeSigma) and 0.9% NaCl (S5886, MilliporeSigma) in high purity water, and then filter-sterilized using a 0.2μM filter (430626, Corning) and ~10mL aliquoted into 15mL tubes (12565268, Fisher Scientifici). Aliquots may be stored at 4°C or frozen at -20°C. This solution should be prepared at least one day prior to the assay.
CRITICAL: 30% glucose must be prepared for the treatment of any mice that develop hypoglycaemia during the insulin tolerance test also.

-

- Syringe: 1mL sterile disposable syringe (309659, Becton Dickinson) – one syringe is required per mouse
- Needle: 27G1/2 sterile needle (305109, Becton Dickinson) – one sterile needle is required per mouse.

CRITICAL: The procedure described here is an intraperitoneal (I.P.) glucose tolerance test. As discussed in the limitations above, some investigators, particularly those investigating incretins, will prefer to perform an oral glucose tolerance; in such a case, an oral gavage needle is required instead.

- Glucose test strips: Bayer Contour Blood Glucose Test Strips (Various) – six glucose test strips per mouse are required; extra glucose test strips in case of mishaps, misreads, or defective test strips is a good idea
- Razor blades: 55411050 (Andwin Scientific via VWR) – one razor blade per mouse cage

- Spreadsheet: One or more datasheets per experimenter, with space for entering the weight, volume of glucose to be administered, blood glucose levels at each of six time points, and room for other notes.
- EDTA Blood Collection Tubes: 20.1288.100 (Sarstedt Inc via Fisher) CRITICAL Required if also collecting blood for insulin or other hormone measurements.
- Insulin ELISA Kit: 90080 (Crystal Chem) CRITICAL Required if also collecting blood for insulin or other hormone measurements.

Insulin tolerance test

- Insulin: Prepared by diluting 22.5uL of sterile 100U/mL Humilin-R (Lily) in 10mL of filter-sterilized (using a 0.2μM filter (430626, Corning)) 0.9% NaCl.

  CRITICAL: Do not use long-acting insulins.

  CRITCIAL: Diluted insulin does not retain activity indefinitely, and must be freshly prepared (within ~2 hrs of use). Do not filter insulin after dilution.

Grid hanging test

- Spreadsheet with a list of mice, with space for entering the weight and the time to fall for up to three attempts.

EQUIPMENT

General equipment

- Marker – A black or other dark color marker capable of temporarily marking the tail of mice, e.g. 52877-310 (VWR).
- Scale – an electronic balance with sensitivity to at least 0.1 g. Some of us use an Ohaus SP-402 Scout Pro Balance (Various), but many brands are available.
- Timer – A lab or kitchen style timer that can count upwards from zero while displaying both minutes and seconds, e.g. 76204-504 (VWR).
- Weighing container – a glass or plastic container of 300-500mL volume for placing mice in while being weighed.

Cardiac function

- Inhaled anaesthesia system with small nose cone to allow easy access to thoracic area with transducer.
- Vevo 770 / 1100 / 2100 / 3100 System with the recommended transducer for the specific system used, e.g. the MX550D (32-55 MHz linear array) for mice with the Vevo 3100 system. Lamp (e.g. 250 Watt Infra-Red warming bulb) to maintain body temperature.
- Gel warmer (available from Visualsonics) or 37°C waterbath to ensure gel is warmed to body temperature.

Cognitive function

- A dedicated room big enough to allocate a 42 x 32 x 31 centimetre arena and for the experimenter to move freely around the cage with access to a computer. The

behavioural room should be in proximity to the place where the animals are housed to limit stress associated with transportation of the mice. A video camera attached to the ceiling, thus providing overhead footage on the very centre of the arena and connected to the computer. Room lights should be indirect or at least located in a way that no reflection appears in the area to be tracked and should preferably be adjustable in brightness, as illumination should be around 150 lux at the floor of the arena.

CRITICAL: One of the most damaging distractions in the behavioural laboratory is unexpected noise, which should be avoided. Similarly, the room environment should be free from interference by external or internal lights or odours. A running air purifier (e.g. Honeywell True HEPA Air Purifier 50250) in the testing room can be used to generate a constant low level of white background noise to help to mask unexpected external noise sources.

- An open wooden or methacrylate box of 42 x 32 x 31 centimetres used as an open-field arena. Although novel object recognition performance has not been directly compared across boxes with different dimensional characteristics, this is unlikely to impact the result of the test. Indeed other mouse behavioural measures in an open field setting have been found to be insensitive to different box dimensions [85].The box should not be covered with sawdust and should be located on the floor or on a table.

- Small plastic objects, e.g. toys or cans. We use objects sized 5 x 5 x 5 cm but a variety of object dimensions can be used. The objects should not have mobile parts and must be easy to attach and remove from the floor of the arena (usually by means of an adhesive tape). They should be located around the centre of the arena, in an opposite and symmetrical way (see below).

- CRITICAL: It should be established empirically for all pairs of objects to be used for the assessment of novel object recognition that they induce indistinguishable levels of exploratory behaviour when encountered for the first time by animals of the specific mouse strain used for experimentation. Objects should be sufficiently distinct for successful discrimination, yet similar enough (e.g. with regards to size, shape, presence of protrusions/intrusions, texture, brightness) to avoid inherent preference biases for one of the objects within the object pair. For further consideration of object feature influences on the performance in novel object recognition tasks, see [86,87].

- A computer connected to an overhead video camera and able to run commercially available animal behaviour tracking software. The software needs to be able to track the animal in the arena, by a multiple body points tracking system, and based on the tracking data, compute time spent, as well as distance travelled in different user-defined sections of the arena, considering the nose point and center point of the animal separately[23]. We use Ethovision XT (Noldus) or Smart video tracking software (Panlab). Video footage should be provided for potential additional analyses. Careful arrangement of the environment of the behaviour room should be observed (especially illumination). This should follow the instructions by the software manufacturer, to assure adequate tracking of the animals' nose point. We recommend the use of automated video tracking. While manual scoring may be necessary in some circumstances, the process is labour-intensive and the exact determination of whether the nose point is in the target area is more subjective[23]. For automated video tracking, it is critical to maintain continuity of nose tracking. This can be easily achieved with careful illumination combined to adequate contrast detection parameters allowed by the software.

Body Composition

- A variety of systems are available such as the EchoMRI-700 system with an A-100 insert antenna, which enables the measurement of animals up to 100g of body weight, and the Bruker's minispec Whole Body Composition Analyzer LF90.

  CRITICAL: While any system based on nuclear magnetic resonance (NMR) can be utilized, dual-energy X-ray absorptiometry (DXA) based body composition analysis is not preferred, as anaesthesia is required, DXA uses a small dose of ionizing radiation, the length of procedure is longer, and NMR is more accurate than DXA for the determination of both fat mass and lean mass[11]. However, if a laboratory only has access to DXA, the investigator will need to balance these factors vs. the need to acquire body composition data.

- An animal holder, sized appropriately to the weight of the animal. A tube sized for mice of approximately 40g (e.g., EchoMRI part 600-E25130R-40) is appropriate for the majority of longitudinal healthspan studies in wild-type mice.

Energy Balance

- A metabolic chamber system equipped to measure food consumption, spontaneous activity, and respiratory gasses (O_2 consumption, CO_2 production). The Oxymax/CLAMS (Comprehensive Lab Animal Monitoring System) manufactured by Columbus Instruments is the most commonly used, but alternative and functionally equivalent systems are also available from other vendors, e.g. TSE Systems or Sable Systems. Other options are available, including measurement of water intake, running wheels (to assess voluntary activity), urine collection, and temperature telemetry. Some systems can also assess respiratory gasses during forced running with the use of a treadmill.

  CRITICAL: The system should be installed in a location where disturbances from animal facility staff and laboratory personnel is minimal, ideally in a room not used for the housing of other animals and not routinely entered by others while the chambers are in use. The system can be installed in an optional environmental chamber, which we highly recommend as it aids in the isolation of the animals from external stimuli. In particular, it permits finer control of temperature and lighting than is typically possible in an animal facility, and places it under computer control.

- Food processor – if the metabolic chamber system has a feeder that dispenses ground food rather than pellets, a commercial food processor to pulverize pelleted feed into a powder is extremely convenient. Many different options are available such as the Hamilton Beach 12-Cup Stack and Snap Food Processor (#70725A, Amazon). Small quantities of food may also be pulverized by hand using a mortar and pestle.

Glucose and insulin tolerance tests

- Glucometer – a human glucose meter capable of utilizing the test strips purchased in reagents such as a Bayer Contour Blood Glucose Meter (Various), but many brands are available

  CRITICAL: Different models of glucometers consistently produce slightly different blood glucose readings. As a consequence, all strips and glucometers used during an assay should be of the same brand and model.

Grid hanging test

- The apparatus used can basically be as simple as a metal cage top from the home cage of the mouse or any other kind of mesh/grid (squares no bigger than 1cm x 1cm) that allows the mouse to grip and hang upside down.

Grip strength

- A Grip Strength Meter – similar and functionally equivalent grip strength meters are available from several vendors, including Columbus Instruments (e.g., 1027CSM, Single Computerized Sensor with Standard Pull Bars for Mice) and Harvard Apparatus (e.g., Bioseb Grip Strength Meter).

Hanging bar

- Hanging bar apparatus - A 50 cm wide, 2 mm thick metallic wire tightly secured between 2 vertical stands approximately 37-50 cm above a layer of bedding.

Rotarod

- A variable speed, accelerating rotarod –e.g. Columbus Instruments (e.g., 0890M, Rotamex-5 4 Lane Rota-Rod for Mice with RS-232 and Software) but similar equipment is also available from a variety of vendors.

PROCEDURE

Methods for assessment of cardiac function

1. Place ultrasound gel in the gel warmer.
2. Turn on the ultrasound machine. Enter animal ID, date and time, and other relevant information, including body weight, age and experimental intervention (if any).

CRITICAL: Use high frequency ultrasound transducer appropriate to the size and weight of the animal; e.g. a 40 MHz transducer for imaging mice less than ~20 grams or a 30 MHz transducer for mice weighing more than ~20 grams.

CRITICAL: In order to maintain normothermia during the imaging session, turn on the warming mechanism on heated platforms, as well as an ancillary warming lamp if the room temperature is cool (e.g., < 72°F/22°C). It is helpful to have a warming lamp and heat pad also focused on the cages used for recovery (with mice recovering on a paper towel in an otherwise empty cage) to prevent hypothermia and accelerate recovery. Turn these devices on at this time as well.

CRITICAL: Ensure that all hardware/tubing for scavenging of isoflurane gas is connected/intact, functional, and turned on (e.g., suction for active scavenging if present).

3. Gently pick up the mouse by the tail and firmly grasp the animal by its nape. Quickly and accurately insert the animal's nose and mouth into the nosecone and begin administration of 2-3% isoflurane. Once anesthetized (10-30 seconds), lay the animal on the platform in supine position making sure the forefeet and hindfeet lay on the ECG sensors of the platform.

CRITICAL: Methods for induction of anaesthesia are a major source of variability when imaging rodents. If you are using an "induction box" prior to placement of the mouse on the platform/nosecone, be sure to standardize the concentration of isoflurane used for induction and use the animal's individual response to anaesthesia (i.e., immediately

upon immobilization and clear ability to be transferred safely to platform) rather than a fixed period of time for induction.

4. Gently secure the animal with adhesive tape on all four limbs and the tail, and lightly also apply adhesive tape to stabilize the head in the nose cone apparatus.
5. Check heart rate (HR). Ensure that the baseline HR in is between 500 to 700 bpm.

CRITICAL: Ensure that the HR does not fall below 450 bpm under any circumstances. Should HR fall below 450 bpm, progressively reduce anaesthesia by 0.1% increments every 15 seconds until HR recovers. If HR does not recover and/or continues to decrease, immediately remove the animal from the nose cone and transfer to a warming pad for recovery. Adjust anaesthesia flow by small increments accordingly (~0.1 % increments every 15 seconds) until a stable state of anaesthesia is reached. While the final level of anaesthesia can vary based on the type and design of the isoflurane delivery/evacuation system used, experienced imaging personnel can successfully acquire data using level of isoflurane as low as 0.8% (with commonly used ranges of 1.0-1.5%). Critically, this should not be a standardized level but tailored to the physiological response of each animal to anaesthesia. Also, this should be an “imaging plane” of anaesthesia (responsive to toe pinch, no cardiorespiratory depression), not a “surgical plane” of anaesthesia (non-responsive to toe pinch, often cardiorespiratory depression).

TROUBLESHOOTING:.

6. Place ECG gel on contact pads, with a narrow strip of tape to attach the back of the fore paws to the contact pads and the pads of hind paws to the contact pads.
7. Insert temperature probe in rectum and tape in place. Document baseline body temperature and ensure that the warming lamps and platform are sufficient to maintain body temperature throughout initial imaging sessions (particularly after shaving).
8. Shave off hair on chest using an electric clipper designed for use with fine hair. Wipe clean chest with a damp paper towel. Clip fur from the top of the sternum down to the xiphoid process then down either side of the rib cage. Shaving lotion may be used if helpful.
9. Cover clipped area with hair removal lotion, leave for 30 seconds to a minute for mice. Remove lotion with warm wet swabs.

CAUTION: Make sure all hair removal lotion is removed as chemical burns will form if any is left.

10. Apply a generous amount of warmed ultrasound gel on the transducer or directly on the animal’s chest.
11. Position the transducer parasternally about 90 degrees perpendicular with the long axis of the heart with image index marker of the transducer pointing posteriorly. While in B-mode, slide the transducer cephalad until the aortic valve (AV) comes into view.

TROUBLESHOOTING:

12. Press the color Doppler control key in the control panel. Apply color Doppler to region of the aortic valve and carefully sweep back and forth across the aortic root to observe for aortic valve regurgitation (mosaic-patterned, high velocity jet in

the ventricle during diastole) or stenosis (mosaic-patterned, high velocity jet during systole). Importantly, neither of these phenotypes generally emerge with chronological aging in mice.

13. Place the transducer in the apical position in the B-mode. Position the transducer so that it is angled towards the head of the mouse Observe the right ventricle (RV), left ventricle (LV), right atrium (RA) and left atrium (LA) on the image display.

From the apical 4-chamber view, bring the mitral valve in focus by reducing the image width.

14. Using the apical 4-chamber view, apply color Doppler to image flow from the left atrium through the mitral valve during diastole. Observe for mitral valve regurgitation.

15. From the apical long axis view, tilt or point the transducer tip with a rocking motion so that the right ventricle is at the center of the image display. Reduce the image width so that only the right ventricle is visible in the image display.

16. Apply colour Doppler in the region of the tricuspid valve. Observe for tricuspid valve regurgitation.

17. Move the transducer in the parasternal position at the level of aortic valve. Observe the pulmonic valve in the parasternal short axis views.

18. Rotate the transducer clockwise to a modified parasternal long axis position. Then tilt the transducer slightly upward to obtain a short axis view of the pulmonic valve.

19. Apply colour Doppler in the region of the pulmonic valve to assess for valvular regurgitation (mosaic-patterned, high velocity jet during diastole) and stenosis (mosaic-patterned, high velocity jet during systole). Importantly, neither of these phenotypes generally emerge with chronological aging in mice

20. Obtain short axis view of the LV in B-mode with the transducer in parasternal short axis position at the level of papillary muscles.

21. From a parasternal short axis view of the left ventricle, press the M-mode button located in the control panel. Using the track ball, position the M-mode cursor at the level of the papillary muscles, and obtain M-mode images.

22. Measure left ventricular cavity dimension at end-diastole where the distance between the anterior wall and posterior wall is largest, and in end-systole where inward motion of the both anterior and posterior walls is maximum

23. Place the transducer in B-mode imaging modality and move the transducer to the apical window (make sure the transducer is angled towards the head of the mouse). Often times, placing the platform so that the animal is in a slight “head down tilt” can facilitate image acquisition from this window.

24. Place sample volume at the tips of the mitral valve leaflets. Derive the isovolumic relaxation time, isovolumic contraction time, left ventricular ejection time and peak mitral inflow velocity from the spectral display of pulsed-wave Doppler velocities across the mitral valve

25. Perform tissue Doppler imaging (TDI) of the mitral annulus in the apical long axis view. Press the TDI control key and place sample volume at the medial aspect of the mitral annulus. Make sure that the sample volume does not encroach on the

mitral leaflets. Keep Doppler sample volume size between 0.21 mm to 0.27 mm. Measure early diastolic velocity (e') of the mitral annulus.

CRITICAL: Review acquired images. Ascertain that all required images were obtained.

26. Remove the animal from anaesthesia and place the animal on an absorbent paper towel (not bedding, which can be aspirated/block airways during recovery). Observe the animal until sternal recumbency is attained. If the anaesthesia is administered appropriately, recovery should occur within 30 to 60 seconds.

Assessment of cardiac function.

CRITICAL: Generally speaking, high resolution echocardiography systems designed for use in small animals have multiple tools that facilitate calculation of cardiac function parameters. Consequently, this section will focus predominantly on critical aspects of *post hoc* analyses that are central to the acquisition of reproducible cardiac function data.

27. *Ejection fraction, systolic function, and ventricular wall thickness from M-mode* (**Figure 2A**). We typically assess ejection fraction from M-mode images (step 21), as the data tend to be more reproducible and less affected by subtle alterations in probe orientation and/or ventricular movement/orientation throughout the cardiac cycle. For assessment of cardiac dimensions and automated calculation of ejection fraction, measure distance between the anterior and posterior ventricular walls at its largest point during end-diastole, and at its smallest point at end systole. Subsequently measure the anterior and posterior wall thicknesses at end systole and end diastole. While it is essential to use the papillary muscles as an anatomical landmark to ensure an appropriate imaging plane (which should be assessed on each M-mode measurement), it is critical to exclude them from any measurements of ventricular dimension or wall thickness. When possible, it is ideal to take 5 or more measurements during consecutive cardiac cycles to ensure accuracy of the data.

28.

a.

Assessment of diastolic function using pulsed-wave Doppler measurements (**Figure 2B, C**). We generally do not assess diastolic function using E/A ratios due to the fusion of these two peaks at physiological heart rates in mice. Our preferred method is to use the E/e' method, which is derived from the blood and tissue Doppler images acquired in steps 24-25. It has also recently been shown that isovolumic relaxation time (IVRT) can serve as an additional tool to detect changes in diastolic function in the diseased hearts of young mouse models of diastolic dysfunction[88]. Given the high degree of context dependence of most models (and the general lack of validation in models of chronological aging to date), we strongly recommend the acquisition of at least two measurements of diastolic function (e.g., E/e' and IVRT) in studies of aging mice. Future work assessing the utility and validity of other variables in detecting subclinical and overt diastolic dysfunction (e.g., peak longitudinal strain rates, etc.) will be essential to advancing the field of biogerontology. For assessment of peak mitral inflow velocity (E), open the apical long-axis view of the mitral valve acquired in step 24. Place the sample volume for the pulsed-wave Doppler at the tips of the

mitral valve leaflets and acquire the spectral display over several cardiac cycles. Measure the peak inflow velocity, and if desired, the isovolumic contraction, isovolumic relaxation, and ejection times for the left ventricle. For assessment of the septal mitral annulus tissue velocity (e'), place the Doppler sample volume (approximately 0.21-0.27 mm) at the septal region of the mitral annulus. Be sure that the sample volume does not encroach on the mitral leaflets. Measure and record the tissue velocity.

Cognitive function

CRITICAL This novel object recognition test is based on the one developed by Ennaceur and Delacour[81] and adapted from the protocol by Leger *et al.*[82]. This assay is of special translational interest given existing implementations for rodents and humans[90]. The test described here uses the recognition and association of salient cues, specifically the comparison between similar but not equal objects, in order to recognize which of the objects is already known and which one is novel. By adjusting the number of cues, the salience of cues, and the time allowed to learn and to memorize the information provided, the test can efficiently be adapted to mice of different ages and genetic background. In the same way, by adequately changing the type, form, and brightness of the objects in the object recognition test, the test can be adapted to be performed more than one time.

1. *Before the assay.* Plan the experiment by considering the sex of animals and the time of day behavioural assays will be undertaken. Identical procedures can be used on both male and female mice[91]. Notably, while it has previously been thought that female performance is sensitive to the phase of the estrous cycle, recent analyses have found that female variability when studied without regard for the estrous phase is actually not any greater than male variability; thus, female mice may be studied without monitoring of the estrous cycle[92-94]. Most laboratories perform behavioural tests during the daytime, which is the resting time for mice. We recommend that behavioural experiments are executed during the daytime, but avoiding the period of time leading up to the end of the light period when the animals start to become more active. The preferred time to perform the behavioural analysis is between 9 am (approximately two-three hours after the end of the dark cycle) and not after approximately 4-5 pm (the two hours leading up the beginning of the dark cycle).

2. Ensure animals have been handled as this decreases stress and anxiety induced by contact with the investigator. The experimenter can handle the animals while moving them from an old to a new cage every day over the course of two-three days. Gently take the animal by its tail (being careful not to suspend by the tail for too long) and place the animal with one hand on the experimenter's arm covered with a waterproof sleeve, or alternatively on a paper towel in your gloved hand, for 15 s. Once all of the animals in one cage have been handled once, repeat in the same order 5 more times. It is also recommended that the animals are labelled by marking the tail.

3. For aged mice, determine visual acuity as the protocol below relies heavily on the vision of animals, and aged animals may suffer from blindness or reduced vision. Consequently, it is mandatory to perform a short visual discrimination test prior to all stages of the longitudinal analysis. Protocols for visual discrimination test for ageing

mice can be found elsewhere; we recommend the efficient visual discrimination test for ageing mice described by Brooks and colleagues[95], which is based on the animal's normal ability to discriminate between bright and matte objects. Animals showing significantly reduced vision compared with the group mean should not be utilized for the remaining stages of the longitudinal analyses.

4. Prepare the arena and objects by cleaning with an ethanol solution (70%). If a methacrylate arena is used, no alcohol should be used, but a soft detergent solution. This step must also be carried out between all trials to avoid animals being influenced by the scents of previously tested mice.

5. Prepare software. Most software requires the experimenter to draw or locate predefined areas into the arena, which will consist of small areas around both novel and known objects. We consider the mice exploring objects when their nose point, but not their center point[23], is within a predefined area with a distance equal to or less than 2 cm from the object[81]. The maximum time for each trial will need to be set (see below). For specific steps, follow the manufacturer instructions.

CRITICAL Ensure that one specific object is not used only as the novel object or as the familiar object. Change the use of each object within a pair, such that each object is equally used as novel object or familiar object within each experimental group. Similarly, one specific side of the box should not contain always the novel or the familiar objects. However, the assigned locations of the objects during training should be maintained during the test sessions.

6. *Habituation and training.* To Habituate the animals, place the animals in the room for at least 30 minutes prior to commencing habituation to the arena. Then allow animals to explore the open-field arena freely individually in the absence of objects for five minutes[23,96]. Note that center occupancy measures in the open field arena during habituation can be used to gauge stress- and anxiety-related behaviours in the mice [97]. Remove the animals and place back in their home cage.
7. The day after habituation, place two identical objects (objects A and A) in the arena. Individually introduce the animals into the open-field arena facing one of the corners and allow the animal to explore the objects freely (for five-fifteen minutes depending on the desired difficulty). Then remove each animal and put back in its home cage.

TROUBLESHOOTING:

CRITICAL: The time spent exploring the objects will greatly impact the efficiency with which the animal is able to recognize new objects subsequently. This influences

the discrimination index (DI, see below for definition), and allows the experimenter to adjust the difficulty of the test according to the age, mouse strain, and model of ageing. In the same way, the time between training and test phases can be adjusted to modify the difficulty of the task. It is highly recommended to establish in a pilot study how normal ageing and/or geroprotectors may affect novel object recognition in the experimenter's specific mouse model. Special attention should be paid to the time between training and test phase, and the time allowed to explore the objects. The shorter the time allowed to explore, and the longer the delay between training and test phases, the more difficult the test. This timing can then be used as a reference, adjusting the difficulty of the test to the specific mouse model and the intervention.

8. *Test phase.*– Introduce each animal into the open-field arena containing a known object (object A) and a novel object (object B) and allow the animal to explore them freely for 10 minutes. Return all animals to their home cage after the test.CRITICAL To assess short-term object recognition memory (STM), perform the test one hour after the training phase. To assess long-term object recognition memory (LTM), perform the test 24 hours after the training phase or repeat the test with a different novel object (object C) 24 hours after the STM test.

TROUBLESHOOTING:

CRITICAL: At different times after the first test phase, new objects can efficiently be included as new unfamiliar objects, thereby permitting a longitudinal analysis of the cognition abilities of the animals. However, the DI will not be exactly comparable to those obtained at the first time, as prior experience with the testing environment (the arena) and the procedural aspects of the task is expected to influence behavioural outcomes.

Metabolic Health ***Body Composition assay***

1. Plug the EchoMRI machine in at least 24 hours prior to the assay in order to charge the magnet.
2. Calibrate the machine following the manufacturer's protocol.

   CRITICAL: The manufacturer advises (and we have noticed) that the presence of large masses of metal near the machine will interfere with (and may prevent) calibration, and may decrease the accuracy of measurements. The machine must be positioned away from large metal objects, including heavy metal tables, chairs, and gas cylinders.
3. Using a scale, weigh each animal.
4. Place awake mice into an appropriately sized holder, sliding the adjustable barrier to minimize the area of movement for each animal.

   CRITICAL: The mouse needs to be snug, but should not be held tightly. The investigator should not attempt this procedure with progeroid mice or aged mice that are visibly frail unless the scientific goals of the project absolutely require it.

5. Measure whole body fat mass, lean mass, and (if desired) water mass using the instrument in accord with the standard operating procedure for the machine. We typically take duplicate one-minute measurements on each animal, omitting the water stage. CRITICAL: While scanning, the investigator should ensure that the sum total fat mass, lean mass and water mass measured is within ~5% or less of the body weight of the animal determined using a scale. If not, the investigator should re-calibrate the machine, ensuring metal objects are not nearby, and repeat the scan; if this does not solve the problem, the machine may need to be serviced.
6. Average the measurements from the duplicate readings, and express lean, fat and water mass data either as an absolute quantity (mass in grams) or as a percentage of total body weight.

Metabolic Health Energy Balance Assay

1. Arrange a room containing only the metabolic chambers in which the mouse can be undisturbed for two to three days. If arranging a room containing only the chambers is not possible, running experiments when no cage changes are schedule to occur may help minimize disturbances.
2. Turn on the system– ideally for an hour – in order to equilibrate to the desired temperature and to remove humidity within the tubing and chambers.
3. As water vapor needs to be removed from the air before it is flowed over the O_2 and CO_2 sensors, check the water vapor removing system, which differs between systems. Some systems may need a desiccant (e.g. Drierite) replenished or a condensation trap emptied.
4. Calibrate the system (following the manufacturer's instructions), using cylinders of gas with known concentrations of O_2 and CO_2.
5. Fill feeders and water bottles.
    -
    -
6. Weigh each animal. If desired, also determine body composition prior to placing mice into the chamber to enable normalization to lean mass instead of body weight.
7. Place one mouse into each chamber, sealing it according to manufacturer's instructions, and entering the mouse ID and weight into the control program's software.

   CRITCAL: If multiple chamber runs are to be combined, do not always load mice of the same group into the same chamber(s) each time. Randomize mice such that genotypes and sexes are evenly distributed. This eliminates any biases resulting from individual chamber differences (e.g. in noise or lighting).
8. Run the chambers, checking the availability of food and water once per day. As mice are being temporarily single housed in the new environment of the metabolic testing chamber, a degree of habituation is required. Typically, discard data collected during the first 24 hours of chamber time to account for this acclimatization.

CRITICAL STEP: When checking on the animals the first day, the data from the habituation period should be visually inspected to ensure proper functioning of the machine. In particular, the O2 in, CO2 in, O2 out, and CO2 out values and the RER should be checked to ensure that gas concentration values are fluctuation as expected and the RER values are physiologically reasonable.

Glucose tolerance test

1. Fast the mice by placing them in new, clean cage with *ad libitum* access to water, but without food. The length of this fast – and the time of day the experiment is performed – depends upon the question being asked. If primarily interested in pancreatic beta cell function, a short daytime fast of 6 hours is sufficient. However, if interested in hepatic glucose metabolism and insulin sensitivity, a longer overnight 16 hour fast should be performed. As the longer fast provides more information, and the short-term fast is often more variable and uninformative as C57BL/6J mice do not have an age-dependent decrease in beta cell function[99], most aging researchers are best served by performing an overnight fast for 16 hours. The 16h fast should start in the evening (approximately 4-5 PM local time), with the glucose tolerance test performed the following morning, approximately 16 hours later (approximately 8-9 AM local time). If a short fast is performed, the fast should start as soon as the light cycle starts (approximately 6 AM local time), with the glucose tolerance test performed 6 hours later. Special challenges are posed by studies involving once-a-day feeding such as calorie restriction (CR). While there is no perfect answer, if an overnight fast is to be performed, the CR animals should be fed prior to lights out (approximately 4pm). As CR animals typically consume their food within 2 hours, both CR and *ad libitum* fed animals will therefore be fasted for approximately 16 hours the following morning.

   CAUTION: Fasting aged mice overnight (as for a glucose tolerance test) can be stressful, and should be avoided in aged mice (for C57BL/6J mice, we define this as male mice over 30 months of age and female mice over 28 months of age). We recommend ceasing routine glucose tolerance tests after approximately 24 months of age.

   CRITICAL: The bedding should not be caloric; we recommend aspen bedding. The use of commercial sterile water gels (e.g. hydrogel) instead of water bottles is discouraged due to the small number of calories contained within the gel.

2. Move the mice to the procedure room for acclimatization at least one hour prior to the start of the glucose tolerance test. During this time period, you can continue to work preparing the room and the mice (steps 3-5 below), but unnecessary noise should be avoided. Weigh each mouse, marking the tail of each animal with a sharpie or similar marker such that each animal can be readily distinguished from its cage mates without picking it up.

3. Calculate the volume of glucose in µL to be administered to each mouse. Using 30% glucose, a dose of 1g/kg of glucose can be rapidly calculated by dividing the body

weight in grams by 30, and then multiplying by 100 (Example: (30 g mouse/30) x 100 = 100µL). Enter this value on the spreadsheet.

4. Prepare one syringe for each mouse. For either an I.P. or oral glucose tolerance test, attach the appropriate needle, load a 1mL syringe with more than the required volume of glucose, and then remove all bubbles and extra volume from the syringe.

5. To begin the glucose tolerance test, pick up the first mouse, and make a transverse nick across the tail using the edge of a fresh razor blade.

6. Immediately collect a fresh drop of blood with a new glucometer test strip.

7. Gripping the mouse firmly, administer the glucose and start the timer, which will count upwards from 0. Return the first mouse to its cage.

8. Proceed to each mouse in sequence, nicking the tail, collecting a fresh drop of blood with a glucometer test strip, and administering glucose as in steps 6-7. For each of these subsequent mice, note the time of injection of glucose for each mouse on the spreadsheet.

   TROUBLESHOOTING:

9. Determine glucose levels in each mouse 15, 30, 45, 60, and 120 minutes after glucose administration. I.e. when the timer reads “15:00”, it is time to take a second reading from the first mouse that was injected.

   CRITICAL: It is important to collect a fresh drop of blood that is not contaminated with blood from the previous time point; this can be readily achieved by wiping the surface of the tail and then gently stroking the tail until a fresh drop is released.

10. At the end of the assay, ensure that bleeding has stopped in the mice. In rare cases, use of a silver nitrate pencil or styptic powder may be necessary to stop the flow of blood.

Insulin tolerance test

1. Fast the mice by placing them in new, clean cage with *ad libitum* access to water, but without food.

   Many aging researchers routinely fast their mice overnight prior to conducting an insulin tolerance test. However, we believe that this is often unnecessary, and that a short fast – long enough to ensure that all of the mice have not fed recently – is usually sufficient to assess insulin tolerance, whilst being less stressful. Also, mice fasted for 4-6 hours have a relatively high fasting blood glucose relative to mice fasted overnight. This widens the dynamic range of the assay, increasing its power, and also is safer, as it lessens the risk of hypoglycaemia.

   CAUTION: C57BL/6J male mice over 30 months of age and C57BL/6J female mice over 28 months of age should not be fasted overnight. We recommend ceasing routine insulin tolerance tests after 24 months of age.

   CRITICAL: The bedding should not be caloric; we recommend aspen bedding. The use of commercial sterile water gels (e.g. hydrogel) instead of water bottles is discouraged due to the small number of calories contained within the gel.

2. Move the mice to the procedure room for acclimatization at least one hour prior to the start of the insulin tolerance assay. During this time period, you may continue to work preparing the room and the mice (steps 3-5 below), but unnecessary noise should be avoided. Weigh each mouse, marking the tail of each animal with a sharpie or similar

marker such that each animal can be readily distinguished from its cage mates without picking it up.

3. Prepared sterile diluted insulin for injection into the mice as described in “Reagents”.

   CRITICAL: The insulin dose must be adjusted empirically for different strains and conditions; for example, while we utilize 0.75U/kg for *ad libitum* fed C57BL/6J mice, 0.5U/kg is sufficient for CR mice fasted overnight.

   CRITICAL: While this is an insulin tolerance test, at least 10 mL of sterile 30% glucose should also be prepared and brought to the procedure room for the emergency treatment of hypoglycaemic mice.

4. Calculate the volume of insulin in µL to be administered to each mouse.

   We typically dilute 22.5 µL of stock insulin in 10mL of 0.9% sterile saline; using this concentration, the volume of insulin to achieve a final dose of 0.75U/kg can be rapidly calculated by dividing the body weight in grams by 30, and then multiplying by 100 (Example: (30 g mouse/30) x 100 = 100µL). Enter this value on the spreadsheet.

5. Prepare one syringe for each mouse, loading a single 1mL syringe with more than the required volume of diluted insulin, and then removing all bubbles and extra volume from the syringe.

6. To begin the insulin tolerance test, pick up the first mouse, and make a transverse nick across the tail using the edge of a fresh razor blade.

7. Immediately collect a fresh drop of blood with a new glucometer test strip.

8. Gripping the mouse firmly, administer the insulin and start the timer, which will count upwards from 0. Return the first mouse to its cage.

   CRITICAL STEP: Insulin that was not correctly injected into the I.P. space will aggregate and not enter the circulation; the experimenter should note any possibly misinjected animals during this step.

9. Proceed to each mouse in sequence, nicking the tail, collecting a fresh drop of blood with a glucometer test strip, and administering insulin as in steps 7-8. For each of these subsequent mice, note the time of injection of insulin for each mouse on the spreadsheet.

   TROUBLESHOOTING:

   Glucose levels should be determined in each mouse 15, 30, 45, 60, and 120 minutes after glucose administration. Thus, when the timer reads “15:00”, it is time to take a second reading from the first mouse that was injected. CRITICAL: It is important to collect a fresh drop of blood that is not contaminated with blood from the previous time point; this can be readily achieved by wiping the surface of the tail and then gently stroking the tail until a fresh drop is released.

   CRITICAL: Continual assessment of the animal’s health is required throughout the course of the assay in order to avoid hypoglycaemia. Mice with a blood glucose level that rapidly drops below 20 mg/dL, that become cool to the touch, or that become unresponsive should be removed from the assay, injected with approximately 1.5 g/kg sterile glucose, and placed in a new cage by themselves with food on the cage bottom, and monitored until recovered.

CRITICAL STEP: If blood glucose levels do not fall in an animal following insulin administration and misinjection is suspected as the cause, I.P. insulin can be administered a second time immediately following the end of the ITT, and the blood glucose tested 15 minutes later. If the animal now responds robustly, indicating the first dose of insulin was likely misinjected, the ITT data for that animal should be discarded.

10. At the end of the assay, ensure that bleeding has stopped in the mice. In rare cases, use of a silver nitrate pencil or styptic powder may be necessary to stop the flow of blood. Any mice that appear unresponsive or weak can be given 1g/kg glucose. Return food to the cage.

    CRITICAL: Place some food pellets on the bottom of the cage so animals can reach easily.

***Grid hanging test to determine* Muscle strength and neuromuscular function**

CRITICAL The cage top test, also called grid hanging test or four limbs hanging test (as opposed to the wire hanging test where only the forelimbs are used) is a quantitative, inexpensive, non-invasive test used to measure the muscular strength and endurance of mice in opposing the gravitational force with their four limbs. In fact, mice will endure to their maximum strength to instinctively avoid falling. The advantage over the two limbs test is that mice cannot prolong hanging time by balancing or using any other behaviour that may bias the outcomes, so results tend to be more consistent and reliable [100]. The test can be used to measure decline in strength with age, if repeated measurement throughout life is conducted. The ability to hang depends on sex, strain, and age [18,20].

1. Before starting the test, it is important to reduce to a minimum the stress of the animal, so it is suggested to avoid changing cage bedding in the previous 24h.
2. Acclimate the mice for one hour to the testing room, and record the bodyweight of the mouse.
3. Train the mice before performing the test for the first time. To do this…., CRITICAL STEP In longitudinal studies, usually training the first time is enough.
4. Place the mouse on the grid, allow him to grip securely, then quickly turn the grid upside down. The grid can be placed onto a holding apparatus or held by hand at a fixed distance from the floor (80 cm) with a soft pad underneath to avoid injuries after the fall.

   CRITICAL: If the mice let go quickly because they lose their fear of falling, the operator may hold the apparatus a bit higher.
5. Start the timer when you invert the grid to record the latency time to fall (in seconds).
6. If the mouse falls before 10 seconds, allow it to repeat the test. If it still falls before 10 seconds on the third trial, record the best latency out of the three.
7. When finished, place the mouse back in the home cage, and test the next mouse. Depending on how the apparatus is set up, it may be possible to test multiple mice at once, especially for young ones that tend to hold on longer. It is not advisable to test multiple old mice, as they tend to fall quickly and it is difficult to properly track all of them. Also, ensure to wipe the grid and the underneath pad with 70% ethanol or chlorine dioxide based sterilant (Clidox) before starting the next cage.

8. Repeat steps 4-7 until each mouse has done the test three times. Allow the mice to rest for 30 minutes in between. CRITICAL: In longitudinal studies, when the mice are still relatively young it is advisable to consider the maximum hanging time (young females can hang for more than 3 hours at 3-6 months) to have a more discriminating power, whilst when comparing the effects of different treatments in mice older than 12 months, usually setting a threshold at 7 minutes is sufficient.
9. Average the three trials and perform a two-way ANOVA and use weight as a covariant.

Grip strength to determine Muscle strength and neuromuscular function

1. As general good practice acclimatize mice to the novel environment for 30-45 minutes prior to testing if the testing is carried out in a different room than their normal housing. This test should be carried out on an open bench, as disruption from the air flow in hoods can affect the results.

   CRITICAL: For longitudinal studies, always perform the test at the same time of day. Mice with missing or injured digits should not be tested.
2. Remove a mouse from the cage, gripping the base of the tail between the thumb and the forefinger with the thumb below the forefinger.
3. Gently lower the mouse over the top of the grid so that only its front paws can grip the grid. Stop when you can feel some pulling tension from the mouse
4. Keeping the torso horizontal, pull the mouse back steadily (not jerking) down the complete length of the grid until the grip is released. Do not allow the rear paws to touch the grid. If the rear paws do touch, then re-do the measurement. If the mouse does not grip the grid properly then repeat the measurement. Similarly, if the mouse turns backwards during the pull, or leaves the bar without resistance.

   CRITICAL STEP: High variability in results can result from inconsistent performance of the procedure by the operator. The operator may hold the mouse at an incorrect angle, and the reading may be influenced by the forces generated by the operator when pulling.
5. Record the value from the meter. Repeat 3 times at 1-minute intervals and select the highest value. Do not repeat the measurement more than 5 times.
6. Front paws alone are generally sufficient, but if time permits, grip strength assessment can be repeated with all four paws on the grid. However, with old mice the risk of fatigue is greater during multiple testing; allowing mice to rest at least 15 minutes before testing all four paws is recommended.
7. Normalize grip strength values by lean mass, or if lean mass is not available, use body weight. Major differences in body weight can influence the results substantially, potentially leading to erroneous conclusions. An alternative approach to analyze these data is to use ANCOVA analysis using body weight or lean mass as a covariate, which is widely accepted as a more statistically rigorous method to address confounding factors such as body weight than simply normalizing to the weight of the animals [101]. As shown in **Fig. 6A**, the senolytic cocktail dasatinib plus quercetin improves the grip strength of mice in this assay.

Hanging bar to determine Muscle strength and neuromuscular function

1. Acclimate animals by placing them on the wire of the hanging bar apparatus for several minutes watching that they don't use their back legs or tail to climb on the wire. Sometimes, the mice are difficult to position such that they use only the front limbs. This is a matter of patiently training the mice by gently repositioning it on the wire in the correct way until they learn. They may need to be repositioned several times because they tend to use only their hind limbs.

   CRITICAL: In longitudinal studies young mice that have not been previously tested need to be trained and acclimated by placing on the wire each day for 5 days. Old mice (24 months and older) usually are not so active, and do not jump on the wire or run on it, and thus can be trained only once prior to testing. The outcome of results will depend on how well they are trained.

   TROUBLESHOOTING:

   CRITICAL STEP: if the mice let go quickly even when they are young, bring the wire to a slightly higher position to induce a fear of falling.

2. On the day of the test, bring all mice to the procedure room 30 min prior to the test to allow animals to acclimatize to local environmental cues.
3. Carefully take mouse at the base of its tail and bring it in proximity to the wire.
4. Let the mouse grasp the wire with the two forepaws only, and slowly lower the hind limbs in such a way that the mouse hangs on the wire only supported by its forepaws.
5. Start the timer as soon as the mouse is released.
6. When a mouse shows improper behaviour (like balancing on or deliberately jumping off the wire, grabs the wire with four paws, or reaches the end of the wire) reposition the mouse on the wire without stopping the timer.
7. If mouse falls before 10 seconds have elapsed, remember what time they fell and quickly repeat the test. If on the 3rd try the mouse still falls before 10 seconds, record the best latency out of 3 attempts.
8. When a mouse falls off the wire, stop the timer and record the hanging time. When mice are able to hang for 5mins take them off the wire and return them to the cage.
9. Record hanging time.
10. Once one trial is completed for all the mice, complete a second trial by repeating steps 3-9. There should be approximately 30 minutes break between trials for each mouse. Repeat steps 3-9 again so each mouse has done a total of 3 trials.
11. Take the best time out of the 3 trials for analysis and normalize results to lean mass or body weight, or perform ANCOVA using weight or lean mass as covariant. As shown in **Fig. 6B**, the senolytic cocktail dasatinib plus quercetin improves the hang time of mice in this assay. In young mice and mice up to 18 month old, test every 6 months. After that, we recommend testing mice monthly.

Rotarod to determine Muscle strength and neuromuscular function

1. CRITICAL Training mice prior to conducting the rotarod assay is essential. An example training schedule is in steps 1 and 2. Place mice on the rotarod at a

constant speed of 4rpm for a minimum of 60s; and up to 300s. If they fall place them back on. Use a timer to count the 60-300s.

2. On the next two dates (Days 2 and 3), train the mice again by placing them on the rotarod at constant speed at 4 rpm for 60-300s.

CAUTION: As it is the case for behavioural experiments, it is important to not make the mice anxious, and thus careful handing is recommended.

CRITICAL STEP If mice have been previously tested (as in a longitudinal study), perform a refresher training run of the mice as above on just the day before testing. Do not train mice on the day of the test in order to avoid fatigue. If the mice fall then place them back on. Nothing needs to be recorded during this refresher.

3. The day after the last day of training bring the mice into the testing room, allowing them to acclimatise to the room for 30 minutes. During this time, weigh all of the mice, marking tails if desired for easier animal identification.
4. Whilst the mice are acclimatising, clean the Rotarod apparatus in accordance with facility protocols and manufacturer's instructions.
5. Put some padding at the bottom to ensure mice are not harmed during the fall. Some of us fill a Ziplock bag with clean aspen bedding; this has the advantage of being sanitizable.
6. Start the rotarod at 4 rpm, and quickly place each mouse in a separate lane of the rotarod; mice should be placed facing away from the experimenter. The maximum number of mice that can be analysed at a single time is determined by the number of lanes on the equipment, but is typically four-five.
7. Once all the mice have settled and are facing away from the investigator, accelerate the rotarod. The acceleration speed and interval can be determined by the investigator; but accelerating from 4 to 40 rpm in 300 sec is typical.
8. The timers for all five lanes will start counting up. If the mouse falls within 10 second from the start put it back on. If it fails three times to remain on the Rotarod as it slowly rotates, remove them from the trial and record the time as 0s.

CRITICAL: if a mouse has not fallen by the end of the 300s, the experiment should be ended and the time recorded as 300 seconds. This maximum time may need to be adjusted in different strains to fully capture the dynamic range of the animals' abilities.

CRITICAL: When the mouse falls, it will usually (but not always) stop the timer. While recording with the computer is a useful timesaver, a paper record should also be kept such that data can be rapidly captured if the time fails to stop.

9. As soon as a mouse falls, carefully remove the mouse from the apparatus and return the mice to the home cage. In the case of passive rotation (the mice hang the rotarod and complete a rotation), reposition the mice once; if this behaviour is repeated, consider this as the time to fall (but make a note).

10. Repeat the test (steps 6-9) so each mouse does it 3 times. CRITICAL: Trials of the same animal should be separated by a minimum of 15 minutes. If comparisons between cohorts analysed on separate days are desired, the resting time should be similar between each cohort.

11. Plot rotarod performance as time to fall, or as maximal speed. A major consideration is that lighter animals, if healthy, typically can stay on the rotarod longer than obese animals simply due to body weight. Rotarod data can be normalized to body weight, although the most rigorous analysis is to perform a 2-way ANCOVA with body weight or lean mass as a covariant. As shown in **Fig. 6C**, the senolytic cocktail dasatinib plus quercetin improves the rotarod performance of mice in this assay.

TIMING

Timing varies each assay in the toolbox, and for most assays in the toolbox timing will increase as the number of animals increase. We have given indicative times for each test for n=10 mice and two operators in table 1. Table 1 also provides a schedule of when to carry out tests over a 6 week period. Conducting any of the above assays requires advance planning!

TROUBLESHOOTING

We have selected these tests for their relative simplicity. Therefore it is unlikely that they do not work. However, variability may be high and differences may not be detected. To address this we have highlighted all the critical steps in each protocol. If any problem is encountered we recommend to carefully re-examine those steps.

Table 2 also gives guidance regarding some of the most commonly encountered problems.

ANTICIPATED RESULTS

Cardiac function

Generally speaking, mice are remarkably resistant to chronic left ventricular pressure overload and have relatively preserved systolic function (but a reasonable hypertrophic response) during all but the most severe insults (e.g., transverse aortic constriction). As a general guideline, we exclude animals that have evidence of significant aortic or mitral valve regurgitation (which is almost always associated with a phenotype of cardiac hypertrophy and dilatation) unless that is a specific aim of the study. Importantly, measures of cardiac mass (either echocardiographic estimates or direct wet weights after euthanasia) should be evaluated as both absolute weights and when normalized by bodyweight.

In our experience, severe systolic cardiac dysfunction is an uncommon phenotype with unstressed chronological aging (i.e., ejection fraction is generally well-preserved and greater than 75% in most strains). One may observe subtle reductions in ejection fraction, however, which are generally associated with increases in end-diastolic operating volumes and larger increases in end-systolic volumes and slight increases in ventricular thickness/mass.

Diastolic dysfunction can be observed with aging in mice, with cardiac hypertrophy and decreased myocardial performance observed during aging[47], and significant changes in the E/e' ratio can become evident with long-term pathophysiological stressors. As is the case with numerous patient populations, increases in the E/e' ratio are indicative of diastolic dysfunction. While our experience to date would suggest that E/e' values ranging from 15-25 are typical in young, unstressed animals, there can be significant variability across strains. With prolonged stressors combined with chronological aging, we have observed values of E/e' that exceed 75 units, and can be associated with increases in ventricular wall thickness and overall mass.

Collectively, these approaches are readily applicable to assessment of cardiac function with aging in small animals/rodents. The procedure described in this toolbox provides a simply way to assess cardiovascular function longitudinally in aging mice. While inbred mouse strains are quite resistant to the development of atherosclerosis[102], mice can and do develop both systolic and diastolic cardiac dysfunction with age[103]. We expect that many geroprotectors will prevent or reverse pre-existing age-dependent cardiac hypertrophy and diastolic dysfunction[35,47,48]. This can be seen by an improvement in left ventricular ejection fraction, improved systolic function (LVIDd) or decreased LVIDs.

Further comprehensive assessments of cardiac and heart valve function (as described previously by Verzosa and colleagues[89]) generally require significant training and experience, and should not be attempted without extensive training to ensure appropriate technique and rigorous/reproducible results.

Cognitive function

The tracking software can output a range of parameters, such as velocity, time spent exploring both objects (novel and known) and time not exploring objects. Tracking software including nose-point detection is more reliable than tracking tools based only on the detection of the center point of the animal[23]. We consider a behaviour as active object exploration if the nose point of the mouse is within the pre-defined target zone with a distance to the object equal to or less than 2 cm[81]. Climbing on the objects itself is not considered object exploration unless it is combined with sniffing the object[82]. When automatically scoring object exploration, climbing behaviours can be excluded by removing data in which the center point of the mouse is located within the target zone[23]. Despite the use of automated systems, it is recommended that an experienced researcher review the output of the software to detect potential misbehaviour of the animals during the test, which cannot be distinguished by the tracking software.

Results should be presented as a Discrimination Index (DI). The DI is the difference between the time exploring the novel and the familiar object, divided by the total amount of exploration of both objects. Exploration is defined as directing the nose to the object at a maximum distance of one centimetre. Some animals may explore very few times, or not explore one of the objects during any of the training or the test phases. Those animals are excluded from the analysis. A positive score means more time spent exploring the novel object and a negative score indicates more time spent exploring the familiar object. A score of 0 indicates no preference for any of the objects. However, a minimum DI threshold of +0.20 was agreed to indicate meaningful novel object discrimination. It is expected that the vast majority of animals will explore the novel object for longer than the known object. An example of data is presented in **Fig.3**.

Geroprotectors such as rapamycin have shown to improve cognitive function with age. Particularly in the context of mouse models of Alzheimer's disease, there is growing

evidence that geroprotectors may be able to treat or prevent the disease and its associated cognitive deficits[52,53,104-107]. We expect that young mice will spend more time exploring new objects as opposed to already known objects. This effect is lost as the mouse ages, and old mice, which will spend equal time with old and new objects. An improvement in this function would be shown by the recovery of this ability to explore new objects for longer.

Body Composition assay of metabolic health

The procedure described in this toolbox is a simple way to measure body composition using magnetic resonance imaging (MRI). Many longevity-promoting interventions are associated with reduced adiposity, including diets in which total calories, protein, or methionine are restricted[12,17,18,108-111]. Some pharmaceutical interventions which extend lifespan, including rapamycin[112] and acarbose[113], are likewise associated with reduced adiposity. On the other hand, senescence cell clearance restores adipose tissue function and prevents fat loss in aged mice[28]. A geroprotector would likely help to preserve and maintain both muscle mass and adipose tissue function, and limit increases in ectopic lipid accumulation during aging.

Energy Balance to assay metabolic health

In this toolbox, we describe a simple procedure for the measurement of energy balance – simultaneous evaluation of food consumption, activity, and energy expenditure as assessed by indirect calorimetry using an open flow respirometry system. Very precise CO_2 and O_2 sensors measure the difference in CO_2 and O_2 concentrations in air volumes flowing through control or animal cages. The amount of oxygen consumed, and the amount of CO_2 produced, over a given period of time can thus be determined.

Specific measurements of interest obtained from the energy balance assay usually include the VCO_2 and VO_2, the Respiratory Exchange Ratio (RER), activity, food consumption, and energy expenditure (heat). While differences between groups can be measured at multiple time points, typically averages during the course of the light and dark cycles are calculated and compared between groups. As shown in **Fig. 4**, RER, food consumption, and energy expenditure are higher at night, as expected since mice are nocturnal, and these values are affected by dietary composition. There are specific issues for each type of data:

- Activity: The Oxymax software for the Columbus Instrument CLAMS system distinguishes between all activity and ambulatory activity; reporting the sum of ambulatory along the X and Y axis may provide the most accurate assessment of activity.
- Energy expenditure: may be normalized to either body weight or lean mass. While many investigators believe normalizing to lean mass is best, this may be inappropriate as it assumes that fat mass has a minimal contribution to metabolism[98]. Analysing both group and body composition effects using analysis of covariance is widely considered to be the gold-standard [98].
- Food consupmtion: It is very important to inspect the food consumption data and remove any data that reports food consumption of less than zero, as this indicates either a scale problem or the mouse “playing” with food.

We expect that the value of the RER will vary between approximately 0.7, which reflects the exclusive utilization of fats, and 1.0, which reflects the exclusive utilization of carbohydrates [3,6,18,114]. In actual fact, the measured RER can temporarily exceed 1.0 in animals as well as humans that are exercising very vigorously or are calorie restricted.

While the effect of geroprotectors on RER and other readout may vary, healthy young animals have increased "metabolic flexibility" – the ability to adapt and respond to changes in metabolic demand[115] – including the switch between fasting and fed states in response to the day-night cycle. Unhealthy conditions, including aging and metabolic disease, impairs metabolic flexibility and the ability of RER to rapidly switch as mice cycle between the dark and light portions of the day. We expect that young and healthier animals – including aged mice treated with effective geroprotectors – will exhibit better metabolic flexibility, and thus a greater difference in RER measured during the dark and light portions of the day.

Glucose tolerance test as an assay of metabolic health

The protocol in this toolbox is an effective way to determine glucose tolerance, which when performed after an overnight fast provides insight into hepatic gluconeogenesis and insulin sensitivity, as well as pancreatic beta cell function. It is expected that glucose will be cleared more rapidly, and hepatic glucose production suppressed more rapidly, if a geroprotector is effective. Calorie restriction, as well as many other geroprotective interventions including protein restriction, metformin, and acarbose improve glucose tolerance in both young and aged mice, reducing fasting blood glucose levels and decreasing area under the curve[3,12,14,17,18,60,62,116]. If important for the analysis, groups of mice can be compared at individual time points using two-way repeated measures ANOVA. However, it is usually more biologically meaningful (https://www.graphpad.com/support/faq/when-does-it-make-sense-to-use-repeated-measures-two-way-anova/), as well as more powerful, to calculate area under the curve (AUC) and use AUC as the basis for the statistical analysis. As shown in **Fig. 5A**, the geroprotector rapamycin impairs glucose tolerance – increasing the AUC – when injected daily, but when dosed intermittently has no significant effect on glucose tolerance.

Insulin tolerance test as an assay of metabolic health

The protocol in this toolbox is an effective way to determine insulin sensitivity, which is often – although not always – increased in interventions that promote metabolic health. As with glucose tolerance, calorie restriction and a number of other geroprotective diets and interventions in mammals are strongly associated with improved insulin sensitivity [3,12,17,61,111,113]. Glucose data for each animal should be normalized to its fasting (basal) blood glucose level (fasting blood glucose = 100% of basal glucose), and plotted as % of basal blood glucose. As with glucose tolerance tests, groups of mice can be compared at individual time points using two-way repeated measures ANOVA. However, it is usually better to calculate the area under the curve (AUC) of the normalized blood glucose values, and use AUC as the basis for the statistical analysis. As shown in **Fig. 5B**, intermittent administration of the geroprotector rapamycin does not impair insulin sensitivity; this is in contrast to daily administration, which does impair insulin sensitivity and increases the AUC [15].

Muscle strength and neuromuscular function

Assessment of muscle and neuromuscular function can be challenging as measurement of these phenotypes may be influenced heavily by the operator, or by how cooperative the individual animal is. In addition severe muscle defects occur at 28-30 months in C57BL6/J mice[117,118]. For this reason improvements are small and difficult to detect. It is advisable that muscle function is tested using multiple measures of outcome to overcome these issues.

Grid hanging test and Grip strength. Grip strength is "a universal measurement used to assess physical competency in older adults" humans[119], and the decline in grip strength with age is predictive for both the risk of inability to perform activities of daily living, an important component of healthspan, and mortality in humans[120]. For this reason it is used in mice[4,12]. Grip strength as measured using a grip strength meter may decline with age in mice, but this has not been observed by every group[119,121]. Interventions which benefit the muscle system should result in an increase in grip strength, and this has been observed by some groups[29,122]. However, this is not always the case.

We believe the variability associated with this assay, and the small size of the effect means that most studies are underpowered. We have observed that minor variations can confound the results, and that different operators even within the same lab can see different results. We have tried to describe the procedure in detail. However, for grip strength it is advisable that new operators are trained by someone experienced.

The grid hanging test, also known as the cage top test, four limbs hanging test, grid wire hang, or inverted cling test, is a widely used and easy to perform assay; grip strength as measured by this assay also declines with age in mice[18,123,124], and we expect it to increase as a result of an geroprotective intervention. This can be a more robust assay than grip strength.

Hanging bar. This test seeks to evaluate motor function measuring forelimb muscle strength of the mouse. It is expressed as latency to fall from the wire. Neuromuscular abnormalities in mice can be detected as well with this test[18,125]. We expect that geroprotective interventions will often increase wire hang time in aged mice[18,19].

Rotarod The rotarod assay is one of the most commonly used tests in the study of aging mouse healthspan. It is measured as the latency (time) to fall. This time declines with age, and neuromuscular abnormalities in mice can be detected as well with this test [44,125,126]. Rotarod running time is positively impacted by interventions including calorie restriction[34], metformin[3], nicotinamide[20], and rapamycin[127], and it is expected that the latency time to fall is increased with the use of geroprotectors[19,122].

TABLE 2: TROUBLESHOOTING

Cardiac function

Step	Problem	Possible reason	Possible Solution
5	No heart rate/ECG signal	Poor contact with platform, lack of signal conduction	Check foot placement and use electroconductive gel if needed, check connections of platform to system
11	Initial poor image quality	Ineffective skin preparation, insufficient ultrasonic gel for contact	Check skin for effective and thorough hair removal, add additional warmed

			gel,
11	Excess LV dilation/poor systolic function with slow, gradual recovery	Excess anesthesia upon induction resulting in anesthesia-induced cardiodepression	Practice rapidly placing the animal on the platform, inducing under the lowest level of isoflurane possible, and rapidly lowering to a minimal plane of anesthesia. Avoid use of anesthesia box for induction when possible (particularly with obese mice)
All steps	Significant bradycardia (HR<450) and poor LV function	Excess administration of anesthesia, excess pressure on chest	Lower anesthesia concentration; ensure negligible pressure is applied to the chest while imaging
All steps	Profound bradycardia (HR<250), LV dilation, and arrhythmias	Adverse response to anesthesia	Remove animal from nose cone and allow to recover fully. Attempt imaging again after full recovery or on another day.

Cognitive function

Step	Problem	Possible reason	Possible Solution
5	Some animals might spend most of the time in a corner, immobile and without exploring	Stress	Special care in the behaviour room environment. If it happens in individual animals only, these should be removed from the experiment. If this is seen in many animals, it may be necessary to identify and eliminate stressor present in the behavioural testing or animal housing environment. Stress related to the experimental procedure can also be attenuated by

			increasing handling and/or habitation time.
5	Some objects promote animals to play with the object including climbing on top of the object and remaining there for a while	This is most often due to mobile pieces in the object or a big surface on top of the object	Pre-select a different pair of objects
6	Lack of increased exploration time of novel vs. familiar object	Stress Insufficient training Lack of interest in objects	Identify and eliminate possible stressors in the behavioural testing or animal housing environment, increase handling and/or habituation time Increase training time Pre-select a different pair of objects

1580

1581 **Metabolic Health**

1582 ***Body Composition***

Step	Problem	Possible reason	Possible Solution
2	Machine unable to calibrate, takes an unusually long time to calibrate, or takes an unusually long time to read a single animal.	Too much metal mass within a 6-foot radius of the MRI.	Increase separation between metal objects and the MRI.
5	Sum of the fat mass, lean mass, and water weight differ by more than 5% from the weight as assessed using a scale.	Too much metal mass within a 6-foot radius of the MRI; alternatively, calibration standard may have leaked.	Increase separation between metal objects and the MRI, replace calibration standard if necessary, and recalibrate the machine.

1583

1584 ***Energy Balance***

Step	Problem	Possible reason	Possible Solution
5	RER results are non-physiological	If the reference cage is only read at the	Set the system to read the reference

	(continuously > 1.0 or < 0.7) and O_2/CO_2 in readings are constant.	beginning of each experimental run, but room air O_2/CO_2 concentrations change, the delta O_2 and delta CO_2 will be calculated incorrectly.	cage (and update O_2 in and CO_2 values) every time it cycles through the experimental chambers. once before each reading
5	RER results are non-physiological (continuously > 1.0 or < 0.7)	Valves on one or more of the calibration air tanks may have been shut during the calibration procedure; alternatively, the calibration air may have settled or the tank may be empty.	Recalibrate system, checking that there is adequate gas flow from the calibration tanks.
5	The gas values read from all of the chambers (O_2 out/CO_2 out) becomes “stuck” on the same value	A bubble may have formed in the ammonia traps.	Replace the ammonia traps.
Data analysis	Food consumption values are negative or absurdly high.	Mice may have greater than intended access to the feeder, and either deposit waste products or remove food without eating it; or the scale may be defective.	Discard data from cages with negative food consumption values, and inspect the cage bottom to determine if food was removed from the feeder and not consumed. The cage floor may be able to be reset to limit the access of small/lean mice. Verify that the scale is functioning properly and replace/repair as needed.

Glucose tolerance test

Step	Problem	Possible reason	Possible Solution
8	Blood glucose reading seems extraordinarily high or low based on reading at prior time point.	Approximately 2% of reads are unusual, due to causes that include defective or contaminated glucose test strips or	Wipe the tail, and with a fresh drop of blood immediately repeat the reading with a fresh test strip. If value differs by

		samples, or samples that are too small or partially clotted.	more than 10% from the first reading, use the data from this second reading; otherwise, keep the first reading.
8	Blood glucose monitor says all or most glucose test strips are defective.	Monitor failure.	Replace blood glucose monitor. As many monitors are inexpensive, spares can often be kept on hand.

Insulin tolerance test

Step	Problem	Possible reason	Possible Solution
9	Blood glucose reading seems extraordinarily high or low based on reading at prior time point.	Approximately 2% of reads are unusual, due to causes that include defective or contaminated glucose test strips or samples, or samples that are too small or partially clotted.	Wipe the tail, and with a fresh drop of blood repeat the reading with a fresh test strip. If value differs by more than 10% from the first reading, use the data from this second reading; otherwise, keep the first reading.
9	Blood glucose monitor says all or most glucose test strips are defective.	Monitor failure.	Replace blood glucose monitor. As many monitors are inexpensive, spares can often be kept on hand.
8 and 10	Blood glucose level does not drop or rises following insulin administration.	Insulin that is not correctly injected into the I.P. space will aggregate and not enter the circulation.	If misinjection is suspected, I.P. insulin can be administered a second time immediately following the end of the ITT, and the blood glucose tested 15 minutes later. If the animal now responds robustly, the ITT data for that animal should be discarded as a result of misinjection.

Muscle strength and neuromuscular function

Grid hanging test

Step	Problem	Possible reason	Possible Solution
4	Mice let go very quickly	Grid not high enough/loss of fear of falling	Place the grid slightly higher

Grip strength

Step	Problem	Possible reason	Possible Solution
6	Very variable readings	Operator holding the mouse at an angle	Ensure the mouse is kept horizontally when pulling and the force applied is proportionate
6	Very variable readings	Claws are too long	Claw length may be affected by activity or environment (e.g. bedding). Change bedding or clip claws

Hanging bar

Step	Problem	Possible reason	Possible Solution
1	Mouse is difficult to position	Insufficient training	Gently reposition the mouse until they have learnt
1	Mouse let go quickly even if they are young	They have lost fear of falling	Position the bar higher

Rotarod

Step	Problem	Possible reason	Possible Solution
7	Mouse change position on the rotarod,	Insufficient training	Gently reposition the mouse.
7	Mice fall from the rotarod too quickly.	Old, frail mice, background noise.	Gently position mouse on the track again. Reduce exposure to possible distractions. Remove extremely frail mice from the assay.

7	Mouse turns around in the middle of the experiment and start walking in the opposite direction	Insufficient training	Gently reposition the mouse until they have learnt
7	Mouse “hugs” the rod and starting looping	Insufficient training	Gently reposition the mouse until they have learnt

CONCLUSIONS

A lack of standardized techniques to assess healthspan in mice has resulted in reduced reproducibility of results and slowed progress in developing geroprotective agents to extend healthy life. The toolbox described here describes a comprehensive set of techniques that can be used to consistently and reproducibly assess the healthspan of mice. We expect that the use of this toolbox will permit investigators to rapidly and robustly identify candidate geroprotectors that may potentially be of clinical use in the treatment of many age-related diseases in humans.

References

1 Bellantuono, I. Find drugs that delay many diseases of old age. *Nature* **554**, 293-295, doi:10.1038/d41586-018-01668-0 (2018).

2 Tchkonia, T. & Kirkland, J. L. Aging, Cell Senescence, and Chronic Disease: Emerging Therapeutic Strategies. *JAMA* **320**, 1319-1320, doi:10.1001/jama.2018.12440 (2018).

3 Martin-Montalvo, A. *et al.* Metformin improves healthspan and lifespan in mice. *Nature communications* **4**, 2192, doi:10.1038/ncomms3192 (2013).

4 Neff, F. *et al.* Rapamycin extends murine lifespan but has limited effects on aging. *J Clin Invest* **123**, 3272-3291, doi:10.1172/JCI67674 (2013).

5 Zhu, Y. *et al.* The Achilles' heel of senescent cells: from transcriptome to senolytic drugs. *Aging Cell* **14**, 644-658, doi:10.1111/acel.12344 (2015).

6 Arriola Apelo, S. I., Pumper, C. P., Baar, E. L., Cummings, N. E. & Lamming, D. W. Intermittent Administration of Rapamycin Extends the Life Span of Female C57BL/6J Mice. *J Gerontol A Biol Sci Med Sci* **71**, 876-881, doi:10.1093/gerona/glw064 (2016).

7 Richardson, A. *et al.* Measures of Healthspan as Indices of Aging in Mice-A Recommendation. *J Gerontol A Biol Sci Med Sci* **71**, 427-430, doi:10.1093/gerona/glv080 (2016).

8 Hamm, R. J., Pike, B. R., O'Dell, D. M., Lyeth, B. G. & Jenkins, L. W. The rotarod test: an evaluation of its effectiveness in assessing motor deficits following traumatic brain injury. *J Neurotrauma* **11**, 187-196, doi:10.1089/neu.1994.11.187 (1994).

9 Wolff, B. S., Raheem, S. A. & Saligan, L. N. Comparing passive measures of fatigue-like behavior in mice. *Scientific reports* **8**, 14238, doi:10.1038/s41598-018-32654-1 (2018).
10 Schafer, M. J. *et al.* Exercise Prevents Diet-Induced Cellular Senescence in Adipose Tissue. *Diabetes* **65**, 1606-1615, doi:10.2337/db15-0291 (2016).
11 Halldorsdottir, S., Carmody, J., Boozer, C. N., Leduc, C. A. & Leibel, R. L. Reproducibility and accuracy of body composition assessments in mice by dual energy x-ray absorptiometry and time domain nuclear magnetic resonance. *Int J Body Compos Res* **7**, 147-154 (2009).
12 Potter, P. K. *et al.* Novel gene function revealed by mouse mutagenesis screens for models of age-related disease. *Nature communications* **7**, 12444, doi:10.1038/ncomms12444 (2016).
13 Xie, K. *et al.* Epigenetic alterations in longevity regulators, reduced life span, and exacerbated aging-related pathology in old father offspring mice. *Proc Natl Acad Sci U S A* **115**, E2348-E2357, doi:10.1073/pnas.1707337115 (2018).
14 Lamming, D. W. *et al.* Young and old genetically heterogeneous HET3 mice on a rapamycin diet are glucose intolerant but insulin sensitive. *Aging Cell* **12**, 712-718, doi:10.1111/acel.12097 (2013).
15 Lamming, D. W. *et al.* Rapamycin-induced insulin resistance is mediated by mTORC2 loss and uncoupled from longevity. *Science* **335**, 1638-1643, doi:10.1126/science.1215135 (2012).
16 Arriola Apelo, S. I. *et al.* Alternative rapamycin treatment regimens mitigate the impact of rapamycin on glucose homeostasis and the immune system. *Aging Cell* **15**, 28-38, doi:10.1111/acel.12405 (2016).
17 Stout, M. B. *et al.* 17alpha-Estradiol Alleviates Age-related Metabolic and Inflammatory Dysfunction in Male Mice Without Inducing Feminization. *J Gerontol A Biol Sci Med Sci* **72**, 3-15, doi:10.1093/gerona/glv309 (2017).
18 Alfaras, I. *et al.* Health benefits of late-onset metformin treatment every other week in mice. *NPJ Aging Mech Dis* **3**, 16, doi:10.1038/s41514-017-0018-7 (2017).
19 Diaz-Ruiz, A. *et al.* Overexpression of CYB5R3 and NQO1, two NAD(+) -producing enzymes, mimics aspects of caloric restriction. *Aging Cell*, e12767, doi:10.1111/acel.12767 (2018).
20 Mitchell, S. J. *et al.* Nicotinamide Improves Aspects of Healthspan, but Not Lifespan, in Mice. *Cell Metab* **27**, 667-676 e664, doi:10.1016/j.cmet.2018.02.001 (2018).
21 Allard, J. S. *et al.* Prolonged metformin treatment leads to reduced transcription of Nrf2 and neurotrophic factors without cognitive impairment in older C57BL/6J mice. *Behav Brain Res* **301**, 1-9, doi:10.1016/j.bbr.2015.12.012 (2016).
22 Lee, S. *et al.* Assessment of Cognitive Impairment in a Mouse Model of High-Fat Diet-Induced Metabolic Stress with Touchscreen-Based Automated Battery System. *Exp Neurobiol* **27**, 277-286, doi:10.5607/en.2018.27.4.277 (2018).
23 Benice, T. S. & Raber, J. Object recognition analysis in mice using nose-point digital video tracking. *J Neurosci Methods* **168**, 422-430, doi:10.1016/j.jneumeth.2007.11.002 (2008).
24 Whitehead, J. C. *et al.* A clinical frailty index in aging mice: comparisons with frailty index data in humans. *J Gerontol A Biol Sci Med Sci* **69**, 621-632, doi:10.1093/gerona/glt136 (2014).
25 Feridooni, H. A., Sun, M. H., Rockwood, K. & Howlett, S. E. Reliability of a Frailty Index Based on the Clinical Assessment of Health Deficits in Male C57BL/6J Mice. *J Gerontol A Biol Sci Med Sci* **70**, 686-693, doi:10.1093/gerona/glu161 (2015).
26 Kane, A. E., Keller, K. M., Heinze-Milne, S., Grandy, S. A. & Howlett, S. E. 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**, 275-282, doi:10.1093/gerona/gly117 (2019).
27 Sukoff Rizzo, S. J. *et al.* Assessing Healthspan and Lifespan Measures in Aging Mice: Optimization of Testing Protocols, Replicability, and Rater Reliability. *Curr Protoc Mouse Biol* **8**, e45, doi:10.1002/cpmo.45 (2018).

28 Palmer, A. K. *et al.* Targeting senescent cells alleviates obesity-induced metabolic dysfunction. *Aging Cell*, e12950, doi:10.1111/acel.12950 (2019).
29 Xu, M. *et al.* Senolytics improve physical function and increase lifespan in old age. *Nat Med* **24**, 1246-1256, doi:10.1038/s41591-018-0092-9 (2018).
30 Mitchell, S. J. *et al.* Effects of Sex, Strain, and Energy Intake on Hallmarks of Aging in Mice. *Cell Metab* **23**, 1093-1112, doi:10.1016/j.cmet.2016.05.027 (2016).
31 Cummings, N. E. *et al.* Restoration of metabolic health by decreased consumption of branched-chain amino acids. *The Journal of physiology* **596**, 623-645, doi:10.1113/JP275075 (2018).
32 Xie, K. *et al.* Every-other-day feeding extends lifespan but fails to delay many symptoms of aging in mice. *Nature communications* **8**, 155, doi:10.1038/s41467-017-00178-3 (2017).
33 McGreevy, K. R. *et al.* Intergenerational transmission of the positive effects of physical exercise on brain and cognition. *Proc Natl Acad Sci U S A*, doi:10.1073/pnas.1816781116 (2019).
34 Lamming, D. W. *et al.* Depletion of Rictor, an essential protein component of mTORC2, decreases male lifespan. *Aging Cell* **13**, 911-917, doi:10.1111/acel.12256 (2014).
35 Zhu, Y. *et al.* New agents that target senescent cells: the flavone, fisetin, and the BCL-XL inhibitors, A1331852 and A1155463. *Aging (Albany NY)* **9**, 955-963, doi:10.18632/aging.101202 (2017).
36 Yousefzadeh, M. J. *et al.* Fisetin is a senotherapeutic that extends health and lifespan. *EBioMedicine* **36**, 18-28, doi:10.1016/j.ebiom.2018.09.015 (2018).
37 Xu, M. *et al.* Targeting senescent cells enhances adipogenesis and metabolic function in old age. *eLife* **4**, e12997, doi:10.7554/eLife.12997 (2015).
38 Mercken, E. M. *et al.* SRT2104 extends survival of male mice on a standard diet and preserves bone and muscle mass. *Aging Cell* **13**, 787-796, doi:10.1111/acel.12220 (2014).
39 Mitchell, S. J. *et al.* The SIRT1 activator SRT1720 extends lifespan and improves health of mice fed a standard diet. *Cell reports* **6**, 836-843, doi:10.1016/j.celrep.2014.01.031 (2014).
40 Minor, R. K. *et al.* SRT1720 improves survival and healthspan of obese mice. *Scientific reports* **1**, 70, doi:10.1038/srep00070 (2011).
41 Kane, A. E. *et al.* Impact of Longevity Interventions on a Validated Mouse Clinical Frailty Index. *J Gerontol A Biol Sci Med Sci* **71**, 333-339, doi:10.1093/gerona/glu315 (2016).
42 Kane, A. E. *et al.* A Comparison of Two Mouse Frailty Assessment Tools. *J Gerontol A Biol Sci Med Sci* **72**, 904-909, doi:10.1093/gerona/glx009 (2017).
43 Yu, D. *et al.* Short-term methionine deprivation improves metabolic health via sexually dimorphic, mTORC1-independent mechanisms. *FASEB J* **32**, 3471-3482, doi:10.1096/fj.201701211R (2018).
44 Ramos, F. J. *et al.* Rapamycin reverses elevated mTORC1 signaling in lamin A/C-deficient mice, rescues cardiac and skeletal muscle function, and extends survival. *Sci Transl Med* **4**, 144ra103, doi:10.1126/scitranslmed.3003802 (2012).
45 Darcy, J. *et al.* Increased environmental temperature normalizes energy metabolism outputs between normal and Ames dwarf mice. *Aging (Albany NY)*, doi:10.18632/aging.101582 (2018).
46 Cohen, D. E., Supinski, A. M., Bonkowski, M. S., Donmez, G. & Guarente, L. P. Neuronal SIRT1 regulates endocrine and behavioral responses to calorie restriction. *Genes Dev* **23**, 2812-2817, doi:10.1101/gad.1839209 (2009).
47 Dai, D. F. *et al.* Altered proteome turnover and remodeling by short-term caloric restriction or rapamycin rejuvenate the aging heart. *Aging Cell* **13**, 529-539, doi:10.1111/acel.12203 (2014).
48 Flynn, J. M. *et al.* Late-life rapamycin treatment reverses age-related heart dysfunction. *Aging Cell* **12**, 851-862, doi:10.1111/acel.12109 (2013).

49 Bitto, A. *et al.* Transient rapamycin treatment can increase lifespan and healthspan in middle-aged mice. *eLife* **5**, doi:10.7554/eLife.16351 (2016).

50 Chen, C., Liu, Y., Liu, Y. & Zheng, P. mTOR regulation and therapeutic rejuvenation of aging hematopoietic stem cells. *Sci Signal* **2**, ra75, doi:10.1126/scisignal.2000559 (2009).

51 Harrison, D. E. *et al.* Rapamycin fed late in life extends lifespan in genetically heterogeneous mice. *Nature* **460**, 392-395, doi:10.1038/nature08221 (2009).

52 Hou, Y. *et al.* NAD(+) supplementation normalizes key Alzheimer's features and DNA damage responses in a new AD mouse model with introduced DNA repair deficiency. *Proc Natl Acad Sci U S A* **115**, E1876-E1885, doi:10.1073/pnas.1718819115 (2018).

53 Spilman, P. *et al.* Inhibition of mTOR by rapamycin abolishes cognitive deficits and reduces amyloid-beta levels in a mouse model of Alzheimer's disease. *PLoS One* **5**, e9979, doi:10.1371/journal.pone.0009979 (2010).

54 Baur, J. A. *et al.* Resveratrol improves health and survival of mice on a high-calorie diet. *Nature* **444**, 337-342, doi:10.1038/nature05354 (2006).

55 Fontana, L., Kennedy, B. K., Longo, V. D., Seals, D. & Melov, S. Medical research: treat ageing. *Nature* **511**, 405-407, doi:10.1038/511405a (2014).

56 Mitchell, S. J. *et al.* Daily Fasting Improves Health and Survival in Male Mice Independent of Diet Composition and Calories. *Cell Metab* **29**, 221-228 e223, doi:10.1016/j.cmet.2018.08.011 (2019).

57 Wilkinson, J. E. *et al.* Rapamycin slows aging in mice. *Aging Cell* **11**, 675-682, doi:10.1111/j.1474-9726.2012.00832.x (2012).

58 Turturro, A. *et al.* Growth curves and survival characteristics of the animals used in the Biomarkers of Aging Program. *J Gerontol A Biol Sci Med Sci* **54**, B492-501 (1999).

59 Dumas, S. N. & Lamming, D. W. Next generation strategies for geroprotection via mTORC1 inhibition. *J Gerontol A Biol Sci Med Sci*, doi:10.1093/gerona/glz056 (2019).

60 Harrison, D. E. *et al.* Acarbose, 17-alpha-estradiol, and nordihydroguaiaretic acid extend mouse lifespan preferentially in males. *Aging Cell* **13**, 273-282, doi:10.1111/acel.12170 (2014).

61 Lamming, D. W. Diminished mTOR signaling: a common mode of action for endocrine longevity factors. *SpringerPlus* **3**, 735, doi:10.1186/2193-1801-3-735 (2014).

62 Strong, R. *et al.* Longer lifespan in male mice treated with a weakly estrogenic agonist, an antioxidant, an alpha-glucosidase inhibitor or a Nrf2-inducer. *Aging Cell* **15**, 872-884, doi:10.1111/acel.12496 (2016).

63 Miller, R. A. *et al.* Rapamycin-mediated lifespan increase in mice is dose and sex dependent and metabolically distinct from dietary restriction. *Aging Cell* **13**, 468-477, doi:10.1111/acel.12194 (2014).

64 Fontana, L. *et al.* Decreased Consumption of Branched-Chain Amino Acids Improves Metabolic Health. *Cell reports* **16**, 520-530, doi:10.1016/j.celrep.2016.05.092 (2016).

65 De Leon, E. R. *et al.* Age-Dependent Protection of Insulin Secretion in Diet Induced Obese Mice. *Scientific reports* **8**, 17814, doi:10.1038/s41598-018-36289-0 (2018).

66 Nadon, N. L., Strong, R., Miller, R. A. & Harrison, D. E. NIA Interventions Testing Program: Investigating Putative Aging Intervention Agents in a Genetically Heterogeneous Mouse Model. *EBioMedicine* **21**, 3-4, doi:10.1016/j.ebiom.2016.11.038 (2017).

67 Yuan, R. *et al.* Genetic coregulation of age of female sexual maturation and lifespan through circulating IGF1 among inbred mouse strains. *Proc Natl Acad Sci U S A* **109**, 8224-8229, doi:10.1073/pnas.1121113109 (2012).

68 Ryan, D. P. *et al.* A paternal methyl donor-rich diet altered cognitive and neural functions in offspring mice. *Molecular psychiatry* **23**, 1345-1355, doi:10.1038/mp.2017.53 (2018).

69 Fernandez-Twinn, D. S., Constancia, M. & Ozanne, S. E. Intergenerational epigenetic inheritance in models of developmental programming of adult disease. *Semin Cell Dev Biol* **43**, 85-95, doi:10.1016/j.semcdb.2015.06.006 (2015).

70 Martin-Gronert, M. S. & Ozanne, S. E. Early life programming of obesity. *Med Wieku Rozwoj* **17**, 7-12 (2013).
71 Menting, M. D. *et al.* Maternal obesity in pregnancy impacts offspring cardiometabolic health: Systematic review and meta-analysis of animal studies. *Obes Rev*, doi:10.1111/obr.12817 (2019).
72 Dearden, L., Bouret, S. G. & Ozanne, S. E. Sex and gender differences in developmental programming of metabolism. *Molecular metabolism* **15**, 8-19, doi:10.1016/j.molmet.2018.04.007 (2018).
73 Butterfield, N. N., Graf, P., Ries, C. R. & MacLeod, B. A. The effect of repeated isoflurane anesthesia on spatial and psychomotor performance in young and aged mice. *Anesth Analg* **98**, 1305-1311, (2004).
74 Xu, H. *et al.* Smaller sized inhaled anesthetics have more potency on senescence-accelerated prone-8 mice compared with senescence-resistant-1 mice. *Journal of Alzheimer's disease : JAD* **39**, 29-34, doi:10.3233/JAD-130902 (2014).
75 Li, X. M. *et al.* Disruption of hippocampal neuregulin 1-ErbB4 signaling contributes to the hippocampus-dependent cognitive impairment induced by isoflurane in aged mice. *Anesthesiology* **121**, 79-88, doi:10.1097/ALN.0000000000000191 (2014).
76 Li, R. L. *et al.* Postoperative impairment of cognitive function in old mice: a possible role for neuroinflammation mediated by HMGB1, S100B, and RAGE. *J Surg Res* **185**, 815-824, doi:10.1016/j.jss.2013.06.043 (2013).
77 Campbell, J. E. & Drucker, D. J. Pharmacology, physiology, and mechanisms of incretin hormone action. *Cell Metab* **17**, 819-837, doi:10.1016/j.cmet.2013.04.008 (2013).
78 Linnemann, A. K. *et al.* Glucagon-Like Peptide-1 Regulates Cholecystokinin Production in beta-Cells to Protect From Apoptosis. *Mol Endocrinol* **29**, 978-987, doi:10.1210/me.2015-1030 (2015).
79 Renner, S., Blutke, A., Streckel, E., Wanke, R. & Wolf, E. Incretin actions and consequences of incretin-based therapies: lessons from complementary animal models. *J Pathol* **238**, 345-358, doi:10.1002/path.4655 (2016).
80 Leiter, E. H., Premdas, F., Harrison, D. E. & Lipson, L. G. Aging and glucose homeostasis in C57BL/6J male mice. *FASEB J* **2**, 2807-2811 (1988).
81 Ennaceur, A. & Delacour, J. A new one-trial test for neurobiological studies of memory in rats. 1: Behavioral data. *Behav Brain Res* **31**, 47-59 (1988).
82 Leger, M. *et al.* Object recognition test in mice. *Nature protocols* **8**, 2531-2537, doi:10.1038/nprot.2013.155 (2013).
83 Servick, K. Mouse microbes may make scientific studies harder to replicate. *Science*, doi:10.1126/science.aah7199 (2016).
84 Cheng, C. J., Gelfond, J. A. L., Strong, R. & Nelson, J. F. Genetically heterogeneous mice exhibit a female survival advantage that is age- and site-specific: Results from a large multi-site study. *Aging Cell*, e12905, doi:10.1111/acel.12905 (2019).
85 Kalueff, A. V., Keisala, T., Minasyan, A., Kuuslahti, M. & Tuohimaa, P. Temporal stability of novelty exploration in mice exposed to different open field tests. *Behav Processes* **72**, 104-112, doi:10.1016/j.beproc.2005.12.011 (2006).
86 Heyser, C. J. & Chemero, A. Novel object exploration in mice: not all objects are created equal. *Behav Processes* **89**, 232-238, doi:10.1016/j.beproc.2011.12.004 (2012).
87 Ennaceur, A. One-trial object recognition in rats and mice: methodological and theoretical issues. *Behav Brain Res* **215**, 244-254, doi:10.1016/j.bbr.2009.12.036 (2010).
88 Schnelle, M. *et al.* Echocardiographic evaluation of diastolic function in mouse models of heart disease. *J Mol Cell Cardiol* **114**, 20-28, doi:10.1016/j.yjmcc.2017.10.006 (2018).
89 Casaclang-Verzosa, G., Enriquez-Sarano, M., Villaraga, H. R. & Miller, J. D. Echocardiographic Approaches and Protocols for Comprehensive Phenotypic Characterization of Valvular Heart Disease in Mice. *Journal of visualized experiments : JoVE*, doi:10.3791/54110 (2017).

90 Raber, J. Novel images and novel locations of familiar images as sensitive translational cognitive tests in humans. *Behav Brain Res* **285**, 53-59, doi:10.1016/j.bbr.2015.01.046 (2015).
91 Bettis, T. & Jacobs, L. F. Sex differences in object recognition are modulated by object similarity. *Behav Brain Res* **233**, 288-292, doi:10.1016/j.bbr.2012.04.028 (2012).
92 Prendergast, B. J., Onishi, K. G. & Zucker, I. Female mice liberated for inclusion in neuroscience and biomedical research. *Neurosci Biobehav Rev* **40**, 1-5, doi:10.1016/j.neubiorev.2014.01.001 (2014).
93 Clayton, J. A. & Collins, F. S. Policy: NIH to balance sex in cell and animal studies. *Nature* **509**, 282-283 (2014).
94 McCullough, L. D., McCarthy, M. M. & de Vries, G. J. NIH policy: status quo is also costly. *Nature* **510**, 340, doi:10.1038/510340b (2014).
95 Brooks, S. P., Pask, T., Jones, L. & Dunnett, S. B. Behavioural profiles of inbred mouse strains used as transgenic backgrounds. II: cognitive tests. *Genes Brain Behav* **4**, 307-317, doi:10.1111/j.1601-183X.2004.00109.x (2005).
96 Frick, K. M. & Gresack, J. E. Sex differences in the behavioral response to spatial and object novelty in adult C57BL/6 mice. *Behav Neurosci* **117**, 1283-1291, doi:10.1037/0735-7044.117.6.1283 (2003).
97 Prut, L. & Belzung, C. The open field as a paradigm to measure the effects of drugs on anxiety-like behaviors: a review. *Eur J Pharmacol* **463**, 3-33 (2003).
98 Tschop, M. H. *et al.* A guide to analysis of mouse energy metabolism. *Nature methods* **9**, 57-63, doi:10.1038/nmeth.1806 (2011).
99 Gregg, T. *et al.* Pancreatic beta-Cells From Mice Offset Age-Associated Mitochondrial Deficiency With Reduced KATP Channel Activity. *Diabetes* **65**, 2700-2710, doi:10.2337/db16-0432 (2016).
100 Klein, S. M. *et al.* Noninvasive in vivo assessment of muscle impairment in the mdx mouse model--a comparison of two common wire hanging methods with two different results. *J Neurosci Methods* **203**, 292-297, doi:10.1016/j.jneumeth.2011.10.001 (2012).
101 Karp, N. A., Segonds-Pichon, A., Gerdin, A. K., Ramirez-Solis, R. & White, J. K. The fallacy of ratio correction to address confounding factors. *Lab Anim* **46**, 245-252, doi:10.1258/la.2012.012003 (2012).
102 (U.S.), N. C. f. H. S. Health, United States, 2016 in *With Chartbook on Long-term Trends in Health.* (National Center for Heath Statistics, 2016).
103 Taffet, G. E., Pham, T. T. & Hartley, C. J. The age-associated alterations in late diastolic function in mice are improved by caloric restriction. *J Gerontol A Biol Sci Med Sci* **52**, B285-290 (1997).
104 Majumder, S. *et al.* Lifelong rapamycin administration ameliorates age-dependent cognitive deficits by reducing IL-1beta and enhancing NMDA signaling. *Aging Cell* **11**, 326-335, doi:10.1111/j.1474-9726.2011.00791.x (2012).
105 Lin, A. L. *et al.* Chronic rapamycin restores brain vascular integrity and function through NO synthase activation and improves memory in symptomatic mice modeling Alzheimer's disease. *J Cereb Blood Flow Metab* **33**, 1412-1421, doi:10.1038/jcbfm.2013.82 (2013).
106 Ozcelik, S. *et al.* Rapamycin attenuates the progression of tau pathology in P301S tau transgenic mice. *PLoS One* **8**, e62459, doi:10.1371/journal.pone.0062459 (2013).
107 Caccamo, A., Majumder, S., Richardson, A., Strong, R. & Oddo, S. Molecular interplay between mammalian target of rapamycin (mTOR), amyloid-beta, and Tau: effects on cognitive impairments. *J Biol Chem* **285**, 13107-13120, doi:10.1074/jbc.M110.100420 (2010).
108 Cummings, N. E. & Lamming, D. W. Regulation of metabolic health and aging by nutrient-sensitive signaling pathways. *Molecular and cellular endocrinology* **455**, 13-22, doi:10.1016/j.mce.2016.11.014 (2017).

109 Lamming, D. W. & Anderson, R. M. Metabolic Effects of Caloric Restriction in *eLS.* (John Wiley & Sons, Ltd, 2014).
110 Malloy, V. L. *et al.* Methionine restriction decreases visceral fat mass and preserves insulin action in aging male Fischer 344 rats independent of energy restriction. *Aging Cell* **5**, 305-314, doi:10.1111/j.1474-9726.2006.00220.x (2006).
111 Green, C. L. & Lamming, D. W. Regulation of metabolic health by essential dietary amino acids. *Mech Ageing Dev*, doi:10.1016/j.mad.2018.07.004 (2018).
112 Fang, Y. *et al.* Duration of rapamycin treatment has differential effects on metabolism in mice. *Cell Metab* **17**, 456-462, doi:10.1016/j.cmet.2013.02.008 (2013).
113 Brewer, R. A., Gibbs, V. K. & Smith, D. L., Jr. Targeting glucose metabolism for healthy aging. *Nutr Healthy Aging* **4**, 31-46, doi:10.3233/NHA-160007 (2016).
114 Lusk, G. ANIMAL CALORIMETRY: Twenty-Fourth Paper. ANALYSIS OF THE OXIDATION OF MIXTURES OF CARBOHYDRATE AND FAT. *Journal of Biological Chemistry* **59**, 41-42 (1924).
115 Goodpaster, B. H. & Sparks, L. M. Metabolic Flexibility in Health and Disease. *Cell Metab* **25**, 1027-1036, doi:10.1016/j.cmet.2017.04.015 (2017).
116 Solon-Biet, S. M. *et al.* Dietary Protein to Carbohydrate Ratio and Caloric Restriction: Comparing Metabolic Outcomes in Mice. *Cell reports* **11**, 1529-1534, doi:10.1016/j.celrep.2015.05.007 (2015).
117 Lee, A. S. *et al.* Aged skeletal muscle retains the ability to fully regenerate functional architecture. *Bioarchitecture* **3**, 25-37, doi:10.4161/bioa.24966 (2013).
118 van Norren, K. *et al.* Behavioural changes are a major contributing factor in the reduction of sarcopenia in caloric-restricted ageing mice. *J Cachexia Sarcopenia Muscle* **6**, 253-268, doi:10.1002/jcsm.12024 (2015).
119 Ge, X. *et al.* Grip strength is potentially an early indicator of age-related decline in mice. *Pathobiol Aging Age Relat Dis* **6**, 32981, doi:10.3402/pba.v6.32981 (2016).
120 Hirsch, C. H., Buzkova, P., Robbins, J. A., Patel, K. V. & Newman, A. B. Predicting late-life disability and death by the rate of decline in physical performance measures. *Age Ageing* **41**, 155-161, doi:10.1093/ageing/afr151 (2012).
121 Bogue, M. A. *et al.* Accessing Data Resources in the Mouse Phenome Database for Genetic Analysis of Murine Life Span and Health Span. *J Gerontol A Biol Sci Med Sci* **71**, 170-177, doi:10.1093/gerona/glu223 (2016).
122 Xu, M. *et al.* JAK inhibition alleviates the cellular senescence-associated secretory phenotype and frailty in old age. *Proc Natl Acad Sci U S A* **112**, E6301-6310, doi:10.1073/pnas.1515386112 (2015).
123 Peterson, R. L., Parkinson, K. C. & Mason, J. B. Manipulation of Ovarian Function Significantly Influenced Sarcopenia in Postreproductive-Age Mice. *J Transplant* **2016**, 4570842, doi:10.1155/2016/4570842 (2016).
124 Graber, T. G., Ferguson-Stegall, L., Kim, J. H. & Thompson, L. V. C57BL/6 neuromuscular healthspan scoring system. *J Gerontol A Biol Sci Med Sci* **68**, 1326-1336, doi:10.1093/gerona/glt032 (2013).
125 Simmons, D. A. *et al.* A small molecule TrkB ligand reduces motor impairment and neuropathology in R6/2 and BACHD mouse models of Huntington's disease. *J Neurosci* **33**, 18712-18727, doi:10.1523/JNEUROSCI.1310-13.2013 (2013).
126 Jeong, H. *et al.* Sirt1 mediates neuroprotection from mutant huntingtin by activation of the TORC1 and CREB transcriptional pathway. *Nat Med* **18**, 159-165, doi:10.1038/nm.2559 (2011).
127 Zhang, Y. *et al.* Rapamycin extends life and health in C57BL/6 mice. *J Gerontol A Biol Sci Med Sci* **69**, 119-130, doi:10.1093/gerona/glt056 (2014).

Data availability

All data shown in this paper are available from the authors upon reasonable request.

Acknowledgments

This article is based upon work from COST Action (BM1402: MouseAGE), supported by COST (European Cooperation in Science and Technology). Part of this work has been funded by the European Union Research and Innovation Program Horizon 2020 (Grant Agreement Number 730879). The Lamming laboratory is supported in part by the NIH (AG051974, AG056771, and AG062328) and by the U.S. Department of Veterans Affairs (I01-BX004031), and this work was supported using facilities and resources from the William S. Middleton Memorial Veterans Hospital. RdC, CDG and IN are supported by the Intramural Research Program of the NIA/NIH. This work does not necessarily represent the official views of the NIH, the Department of Veterans Affairs or the United States Government.

Author contributions

Ilaria Bellantuono, Rafael de Cabo, Dan Ehninger, Clara Di Germanio, Ignacio Navas-Enamorado, Sarah Mitchell, Paul Potter, Tamar Tchkonia, Jose Luis Trejo and Dudley Lamming contributed equally to the development of the protocols. Ilaria Bellantuono and Dudley Lamming wrote the manuscript and all the other authors edited and approved.

Competing interests

The authors declare no competing financial interests.

Figure legends

Figure 1. A graphic representation of the experimental design for healthspan assessment in response to a geroprotector of choice in C57BL6/J using the recommended toolbox. At each time point the multiple tests are performed over a period of 4-6 weeks, leaving those which may impact more on the animal at the end. In this case we recommend the following order, novel object recognition, frailty index, grip strength, cage top and hanging bar followed by rotarod, body composition, energy expenditure, glucose tolerance test (GTT), Insulin tolerance test (ITT) and echocardiography. Mo, Month

Figure 2. Representative images from high resolution echocardiography systems. A) Representative image of M-mode image used for assessment of cardiac systolic function (AW = anterior wall, PW = posterior wall, LVEDD = left ventricular diastolic dimensions, LVESD = left ventricular systolic dimensions; arrows denote the thickness/dimension of interest for each variable). **B)** Representative image of pulsed-wave Doppler assessment of peak mitral valve inflow (E) using color Doppler guidance (IVRT = isovolumic relaxation time, IVCT = isovolumic contraction time, LVET = left ventricular ejection time). **C)** Representative image of tissue Doppler assessment of the septal mitral annulus velocity (e') using the spectral velocity display. White arrows denote the point of measurement. Images are reproduced from Verzosa and colleagues[89] with permission. Animal procedures were performed following approval by the Mayo Clinic Institutional Animal Care and Use Committee.

Figure 3. NOR memory enhancement in exercised fathers and their offspring. (A–B) NOR memory enhancement in exercised (RUN) vs. sedentary (SED) animals. Results represent as Discrimination index (DI). A DI score of >0.20 was set ad hoc to determine proper novel-object discrimination; statistically significant within-group differences were also considered between the training phase and the test phases. (A) Before exercising, mice were unable to discriminate the novel object in a difficult protocol. B) After 6 wk of physical exercise, exercised animals showed memory enhancement. For comparisons between independent groups, *P < 0.05; for intragroup differences (LTM vs training), +P < 0.05; tendencies 0.05 > #P < 0.09. Extreme values were removed from the analysis. SED A, n = 4; RUN A, n = 5. Figure adapted with permission from McGreevy et al, 2019, PNAS [33]. The animal procedures were performed following approval by the Committee of Ethics and Animal Experimentation of the Cajal Institute, Ethics Committee of the CSIC, and the Animal Protection Area of the Ministry of Environment of the Community of Madrid. LTM, Long term memory, STM short term memory

Figure 4. Effects of dietary protein level on energy balance. Metabolic chambers were used to assess A) Food Consumption, B) Spontaneous activity (as assessed by wheel running), C-E) Respiration and F) Energy Expenditure over a 24-hour period after approximately 8 weeks on the indicated diets (n=8-9/group, two-tailed t-test, * = p < 0.05). Error bars represent standard error. Figure is adapted from [64] with permission. Animal procedures were performed following approval by the Institutional Animal Care and Use Committee of the William S. Middleton Memorial Veterans Hospital.

Figure 5. Glucose and insulin tolerance tests. A) Glucose tolerance test conducted on male C57BL/6J mice treated with either vehicle or with 2 mg/kg rapamycin (1x/day or 1x/5 days) for 2 weeks (n = 9 per treatment; for GTT, Tukey–Kramer test following two-way repeated measures ANOVA, a = P < 0.05 vehicle vs rapamycin 1x/day, b = P < 0.05 vehicle vs. rapamycin 1x/5 days; for AUC, means with the same letter are not significantly different from each other, Tukey–Kramer test following one-way ANOVA, P < 0.05). Figure is adapted from [16] with permission. B) Insulin tolerance test on female C57BL/6J.Nia mice treated with either vehicle or rapamycin (2 mg/kg) once every 5 days for 8 weeks (n = 7–10 vehicle, n =

10 rapamycin, two-tailed t-test). Test was performed 5 days after the last administration of either vehicle or rapamycin, at the conclusion of an overnight fast. Figure is adapted from [6]with permission. Error bars represent standard error. Animal procedures were performed following approval by the Institutional Animal Care and Use Committee of the William S. Middleton Memorial Veterans Hospital.

Figure 6. Assessment of muscle strength and neuromuscular function. Physical function measurements in 20-month-old male mice treated with dasatinib and Quercetin (D+Q) once every 2 weeks (bi-weekly) for 4 months. (a) Grip strength, (b) hanging endurance, and (c) maximal walking speed on rotarod (relative to baseline) 4 months after drug initiation (*n* = 20 for D+Q; *n* = 13 for V). Results are shown as box and whiskers plots, where a box extends from the 25th to 75th percentile with the median shown as a line in the middle, and whiskers indicate smallest and largest values. *P <0.05; n.s., not significant; Two-tailed Student's t-tests. Figure is adapted from [29]with permission. Animal procedures were performed following approval by the Mayo Clinic Institutional Animal Care and Use Committee.

2031 **Table 1**

2032 Estimated timing for each test for n=10 animals and n=2 operators.

2033

2034

Week #	Test	Comments
1	**NOVEL OBJECT RECOGNITION** • **Day 1.** Habituation: 30 min/mouse • **Day 2.** Training phase: 5-15 min/mouse Test phase: For STM, after 1 hour of training phase. 10 min/mouse. • **Day 3.** Test phase: For LTM, after 24h of training phase. 10min/mouse.	For Day 1, estimated time is 5-6 hours. For Day 2, estimated time is between 3:30h and 5:15h with STM same day or 2:00h - 3h without STM. For Day 3, estimated time is 2h.
2	**MUSCLE STRENGTH AND NEUROMUSCULAR FUNCTION** • **Day 1.** Grid Hanging. 3 trials per animal allowing 30 min in between trials. * • **Day 2.** Grip Strength. 3 trials per animal allowing 30 min in between trials. Time estimate is 3 min per animal. • **Day 2.** Frailty index. 15min/animal • **Day 3.** REST Day • **Day 4.** Hanging Bar. 3 trials per animal allowing 30 min in between trials. *** • **Day 5.** Rotarod. 3 trials per animal allowing 30 min in between trials. Time estimate is 3 min per animal.	*depending the age/sex of the animals, estimated total time might vary between 1h and 4h. ** This will take 2-3h *** depending the age/sex of the animals, estimated total time might vary between 1h and 3h. SIDE NOTE: in some tests, total time estimated will depend of the number of different set-ups available. Some apparatus allow the operator to run multiple animals at a time like multi-channel rotarod or grid hanging test. However the operator must be continually scanning and monitoring all animals during this time.
3	**BODY COMPOSITION and ENERGY BALANCE** • **Day 1.** Body Composition Test. Experiment takes 2 min/mouse. Estimated time is 1h. • **Day 2.** Energy balance test. Prepare the animals, calibrating and setting up the experiment and run the experiment. Estimated time 2h. • **Day 3.** Energy balance test. The software is up and running recording in vivo metabolic data. * • **Day 4.** Energy balance test. The software is up and running recording in vivo metabolic data. * • **Day 5.** Energy balance test. Stop the experiment, exporting the data, cleaning the equipment and putting animals back in home	*schedule based on a 72h experiment set-up with two light cycles and two dark cycles of 12h/each. Side note: During these recording days, animals have to be checked by the operator to make sure they are in good condition and have access to food and water. Also software needs to be checked in order to make sure data is being collected accurately. Two people can expedite the set up of these tests. One person does the body composition measurement, the

	cages and room. Estimated time 2h.	other sets up the *in vivo* metabolic machine for calibration and measurement.
4	**GLUCOSE TOLERANCE TEST** • **Day 0.** Fast animals 6 or 16 hours before the start of the test • **Day 1.** Test day. The last of the time points is at 2 hours after the beginning of the test. To allow for setup and clean up, an additional 1.5-2h is needed. Therefore, total time required is 3.5-4h.	Due to the invasive and stressful character of the test, we let the animals rest the whole week after the test. A skilled person can inject between 20-30 mice within 15min. After the conclusion of the test, animals are returned to their home cage. Mice should be checked after 4-6h to ensure they have returned to normal.
5	**INSULIN TOLERANCE TEST** • **Day 1.** Fast animals. Animals need to be fasted 3-6h before the experiment.* Test. The last of the time points is at 2 hours after the beginning of the test. To allow for setup and clean up, an additional 1.5-2h is needed. Therefore total time required is 3.5-4h.	*Fasting periods might vary between different studies. We propose for longitudinal studies a fasting between 3h and 6h. Due to the invasive and stressful nature of this test, we let the animals rest the whole week after the test. A skilled person can inject between 20-30 mice within 15min. After the conclusion of the test, animals are returned to their home cage. Mice should be checked after 4-6h to ensure they have returned to normal.
6	**CARDIAC FUNCTION** **Day 1.** Echocardiogram. Mice are anesthetized using isoflurane-oxygen Echocardiograms can take up to 30-45min per animal depending on the age, strain and bodyweight. Obese mice can be harder to image given their excess amounts of fat. **Day 2.** Data processing. 1h	Caution must be taken for old mice who are more vulnerable to death from isoflurane. A skilled operator can process up to 10 animals per day. Mice should be given at least 5 days (longer for aged mice) to recover from the isoflurane.

2035