"Measurements of damage and repair in aging mice and humans reveals that robustness and resilience decrease with age, operate over broad timescales, and are affected differently by interventions"

S Farrell[1*], AE Kane[2], E Bisset[3], SE Howlett[3,4], AD Rutenberg[1*]

[1]Department of Physics and Atmospheric Science, Dalhousie University, Halifax, Nova Scotia, Canada

[2]Blavatnik Institute, Department of Genetics, Paul F. Glenn Center for Biology of Aging Research at Harvard Medical School, Boston, Massachusetts, USA

[3]Department of Pharmacology, Dalhousie University, Halifax, Nova Scotia, Canada

[4]Department of Medicine (Geriatric Medicine), Dalhousie University, Halifax, Nova Scotia, Canada

[*]**Corresponding authors:**

Spencer Farrell, Andrew Rutenberg

Department of Physics and Atmospheric Science, Dalhousie University,
Halifax, Nova Scotia, Canada B3H 4R2
902-414-8206

spencer.farrell@dal.ca, adr@dal.ca

Abstract

As an organism ages, its health-state is determined by a balance between the processes of damage and repair. Measuring these processes requires longitudinal data. We extract damage and repair rates from repeated observations of health deficits in mice and humans to explore the contribution of robustness and resilience, which respectively decrease damage rates and increase repair rates, to aging health. We find a conserved decline with age in robustness and resilience in mice and humans, implying that both contribute to worsening health in aging. A decline in robustness, however, has a greater effect than a decline in resilience on accelerating damage accumulation with age, and a greater association with reduced survival. We also find that deficits are damaged and repaired over a wide range of timescales ranging from the shortest measurement scales towards organismal lifetime timescales. We explore the effect of systemic interventions that have been shown to improve health, including the angiotensin-converting enzyme inhibitor enalapril and voluntary exercise for mice, and household wealth for humans. We find that these interventions affect both damage and repair rates – but in different proportions for different interventions. These findings have implications for how health in aging, and interventions targeting health, are conceptualized and assessed.

Introduction

As organisms age, they can be described by a health state that evolves according to dynamical processes of damage and repair. The health state is the net result of accumulated damage and consequent repair (Howlett and Rockwood, 2013). Studies of aging have mostly focused on the health-state rather than the underlying dynamic processes, due the difficulty of their measurement. Two common approaches to measuring individual health-states, the Frailty Index (FI) (Mitnitski et al., 2001) and the Frailty Phenotype (Fried et al., 2001), are assembled from heath state data at a specific age and do not separate dynamic damage and repair processes. Nevertheless, strong associations between frailty measures and adverse health outcomes (Hoogendijk et al., 2019; Howlett et al., 2021) indicate that frailty has a strong effect on underlying dynamical processes. This is supported by the increasing rate of net accumulation of health deficits with worsening health (Mitnitski et al., 2007; Kojima et al., 2019).

Reduced resilience, or the decreasing ability to repair damage (or recover from stressors), is increasingly seen as a key manifestation of organismal aging (Ukraintseva et al., 2021; Kirkland et al., 2016; Hadley et al., 2017; Gijzel et al., 2017). Resilience is often assessed by the ability to repair following an acute stressor, such as a heat/cold shock, viral infection, or anesthesia; or a non-specific stressor such as a stochastic fluctuation of the health state, typically within a short timeframe (Scheffer et al., 2018; Gijzel et al., 2019; Rector et al., 2021; Arbeev et al., 2019; Colón-Emeric et al., 2020; Pyrkov et al., 2021). Robustness, or an organism's resistance to damage, has not been as well studied – but there is also evidence for its decline with age (Arbeev et al., 2019; Kriete, 2013). Both resilience and robustness sustain organismal health during aging, but their relative importance and their timescales of action remain largely unexplored.

To study resilience and robustness in aging, we have here developed a novel method of analysis that uses longitudinal binarized health-deficit data from mice and humans to obtain summary measures of organismal damage and repair processes over time. This approach can be adapted to use any biomarker, and is not restricted to biomarkers specifically associated with resilience (Ukraintseva et al., 2021). We apply our method to study how resilience and robustness evolve with age and how they differ between species, between sexes, and under different health interventions.

Developing interventions to extend lifespan and healthspan is the goal of geroscience (Kennedy et al., 2014; Sierra, 2016; Sierra et al., 2021). While some interventions that affect aging health have been identified, how they differentially affect damage and repair, and their timescales of action, is less understood. We consider interventions in mice that have previously been shown to have a positive impact on frailty, the angiotensin converting enzyme (ACE) inhibitor enalapril (Keller et al., 2019) and voluntary exercise (Bisset et al., 2021). In humans, we stratify individuals within the English Longitudinal Study of Aging by net household wealth (Phelps et al., 2020). Wealth is a socioeconomic factor associated with aging health (Zimmer et al., 2021). Understanding how various interventions affect aging health by affecting resilience and

robustness will better enable us to improve and combine interventions to fulfill the geroscience agenda.

Results

Measuring resilience and robustness

A well-established approach to quantify health in both humans and in animal models is to count binarized health deficits in a Frailty Index (FI) (Mitnitski et al., 2001; Whitehead et al., 2014). In longitudinal studies, the FI can be assessed at each follow-up. Here we use longitudinal binarized FI data from mice and humans to quantify organismal damage and repair processes over time. As illustrated in the schematic in Figure 1, the change in number of deficits from one follow-up to the next is determined by the number of new deficits (indicating damage, with deficit values transitioning from 0 to 1, red arrow) minus the number of repaired deficits that were previously in a damaged state (with transitions of deficit values from 1 to 0, green arrow). These counts of damaged and repaired deficits between follow-ups represent summary measures of the underlying damage and repair processes. We model this process with a Bayesian Poisson model for counts of damaged and repaired deficits, using age-dependent damage and repair rates. This is a joint longitudinal-survival model, which couples the damage and repair rates together with mortality.

In our approach, damage rates are the probability of acquiring a new deficit per unit of time, and repair rates are the probability of repairing a deficit per unit time. These are measures of susceptibility to damage (lack of robustness), and ability to repair (resilience). Since the FI is a whole organism-level summary measure of health, these damage and repair rates are also whole organism-level measures of robustness and resilience.

Both resilience and robustness decline in aging populations

We first establish the trends of repair and damage rates in aging. In Figure 2, we plot the age-dependence of the repair and damage processes in mice and humans for three mouse datasets a) 1 (Keller et al., 2019); b) 2 (Bisset et al., 2021); and c) 3 (Schultz et al., 2020); and d) humans from the ELSA dataset (Phelps et al., 2020). Humans are plotted by decade of baseline age at entry to the study to separate out recruitment effects. Points are binned averages from the data, and lines are posterior samples from the model of the rates. Posterior predictive checks show good model quality, seen in Figure 2-figure supplement 1 for mice (a-c) and humans (d).

In each of these datasets, there is a strong decrease in repair rates and increase in damage rates with age (except in mouse dataset 2 for damage rates). Spearman rank correlations ρ for each plot are also shown in Figure 2, highlighting the increase or decrease in rates with age, and 95% posterior credible intervals of these correlations are shown in brackets. Overall, we observe decreasing repair rates and increasing damage rates with age which signify decreasing resilience and robustness with age in both mice and humans. Decreasing repair and increasing damage both contribute to an increasing FI with age in mice and humans (shown in Figure 2-figure supplement 2a-d). We also observe higher FI scores in females versus males in both mice and humans, as reported previously (Kane et al., 2019; Gordon and Hubbard, 2020; Kane and Howlett, 2021).

We evaluate the contributions of damage and repair rates to survival using a joint longitudinal-survival model in mice. In Figure 2e-g), we show that damage rates have much larger hazard ratios for death than repair rates. These hazard ratios are for a fixed FI, itself a strong predictor of mortality in mice and in people (Rockwood et al., 2017), which shows that an increasing susceptibility to damage leads to larger decreases in survival than a comparable decline in resilience. Individuals survive longer when damage is avoided altogether, as compared to damage that is subsequently repaired. This intuitive result indicates that there may be lingering effects of the original damage and suggests that interventions that focus on robustness may be more effective than those that focus on resilience.

The acceleration of damage accumulation is determined by a decline in robustness

The plots of FI vs. age shown in Figure 2-figure supplement 2 (see also Mitnitski et al., 2001; 2005; 2012; 2013) has a positive curvature, accelerating upwards near death (Stolz et al., 2021). This positive curvature is also seen in other summary measures such as Physiological Dysregulation (Arbeev et al., 2019). However, the origin of this curvature is unknown -- whether it is due to a late-life decrease in resilience or a decline in robustness.

We measure the curvature of the FI with the second time-derivative, which can be computed with the age-slopes of the damage and repair rates (see Methods). In Figure 3, we show the separate contributions to this curvature, separated into terms involving damage (pink) and terms involving repair (green). Summing these terms, we observe the typical positive curvature that indicates an acceleration of damage accumulation.

We find that the decline in robustness (indicated by the damage rate terms) has the strongest effect on the curvature of the FI. In mice, this is seen in every dataset and is significant at the indicated ages (Figure 3b-d) and for humans at older ages (Figure 3e). This observed effect indicates that it is the increase in susceptibility to damage with age, rather than the decline of repair, that causes this acceleration of damage accumulation. Credible intervals evaluating the significance of these effects are shown in Figure 3-figure supplement 1.

In Figure 2 (e-g) we had shown that the decline in robustness has the strongest effect on survival in mice. Together with the results shown in Figure 3, these results highlight the important role of declining robustness in aging – in addition to declining resilience (Ukraintseva et al., 2021).

Interventions modify damage and repair rates in mice and humans

Mouse datasets 1 (Keller et al., 2019) and 2 (Bisset et al., 2021) have additional intervention groups treated with either the ACE inhibitor enalapril, or voluntary aerobic exercise, respectively. In Figure 4-figure supplement 1a) and b), we show that these interventions target both repair and damage processes, resulting in lower FI damage accumulation over time for the treated groups. In Figure 4 a) and b), we investigate the effects of these interventions on the curvature of the FI. This curvature is strongly reduced by exercise in mouse dataset 2, with a weaker effect for enalapril (the credible intervals of the intervention effects are shown in Figure 3-figure supplement 1d and e). Notably, exercise stops the acceleration in damage accumulation in both male and female mice by reducing the curvature to zero.

The effect of these interventions on the repair and damage rates is seen in Figure 4c) and d), where 95% credible intervals for the age-slopes show the rate of increase or decrease of the repair and damage rates as age increases. These slopes include both the change in the rate with age, and the effect due to increasing FI with age. Interventions affect the rate of decrease of both repair and damage rates with time, resulting in less cumulative damage.

As shown in Figure 4c, enalapril attenuates the rate of decrease of repair rates in both male and female mice, resulting in age-slopes closer to zero than for controls. Significance is shown with asterisks (*) at the 0.05 level. In Figure 4-figure supplement 1 we show a significant reduction in damage rate (but not slope) for male and female mice with enalapril. A sex-specific effect is seen for voluntary exercise. For female mice, voluntary exercise leads to stoppage of the decline in repair rates (to an approximately zero slope), whereas for male mice it just attenuates the decline (Figure 4d). For damage rates, female mice exhibit an attenuation of the rise with age whereas in male mice exercise stops the age-dependent rise exhibited by control mice.

For humans, we use net household wealth as a socioeconomic environmental factor that serves as a proxy for medical and behavioural interventions that are not individually tracked. As such we report correlations of wealth with repair and damage rates with age, rather than age-slopes after a specific intervention is initiated. In Figure 4-figure supplement 2, we show rates stratified by

terciles of net household wealth, where the lowest tercile exhibits lower repair rates and higher damage rates for younger ages. Correspondingly, the FI is lower for individuals with a higher net household wealth. Treating the wealth variable as continuous, Figure 4e) shows that repair rates are positively correlated with net household wealth, while damage rates are negatively correlated – with significant and stronger effects at younger ages. These results reinforce the findings in mice, where interventions impact both damage and repair rates. In humans, we also see some evidence of decreasing effects of wealth with age – though these may be confounded by recruitment effects of baseline age.

Damage and repair have broad timescales

In the results above, we considered the average damage and repair transition rates vs age. Since individual deficits undergo stochastic transitions between damaged and repaired states, we can also measure the lifetime of these individual deficit states (see Figure 5a). These lifetimes are interval censored (transitions typically occur between observation times) and can be right-censored (death or drop-out before transition occurs). We use an interval censored-analogue to the standard Kaplan-Meier estimator for right censored data (see Methods) to estimate state-survival curves of individual damaged or repaired states. These state-survival curves in Figure 5, considering all possible deficits, represent the probability of a deficit remaining undamaged vs time since a repair transition, or remaining damaged vs time since a damage transition.

We generally observe a significant drop of state-survival probability at early times, indicating some rapid state transitions at or below the interval between measurements. However, all of the curves also extend to very long times – towards the scale of organismal lifetime – indicating that both robustness and resilience operate over a broad range of timescales. These results highlight that repair can occur a long time after damage originally occurred. Note that the timescale of robustness as measured here is not robustness after a specific stressor, but robustness due to the continual stressors of aging. A similar form of non-specific robustness has been measured in a previous study, using the onset age of disease (Arbeev et al., 2019).

As shown by exponential time-scales of resilience and robustness for individual deficits in Figure 5-figure supplements 1 and 2, mice deficits and human deficits exhibit a variety of time-scales of resilience and robustness. Some deficits repair soon after damage (or damage soon after repair), and some repair (or damage) over a broad range of time-scales. The combination of all of these deficits result in the shape of the state-survival curves in Figure 5.

For the interventions studied, there are no dramatic changes of resilience or robustness timescales exhibited in mice. Exercise in the male mice slightly shifted the timescale of resilience, such that deficits were repaired faster in mice that were exercised compared to controls. We expect that we would observed stronger effects of the interventions on these time-scales if we had sufficient data to resolve the impact of the time at which the initial damage or repair event occurred – here we have grouped all times together. For humans (see Figure 5f), we see strong and significant effects on resilience and robustness timescales from household wealth in females, but not males. These effects are particularly strong for the damage timescales that characterize robustness: states remain healthy longer at higher wealth terciles.

Discussion

We have presented a new approach for the assessment of damage (robustness) and repair (resilience) rates in longitudinal aging studies with binarized health deficits. With this approach, we have shown that both humans and mice exhibit increasing damage and decreasing repair rates with age, corresponding to decreasing robustness and resilience respectively. We also demonstrate that decreasing robustness and resilience with age contribute to the acceleration of aging for organisms. Decreasing robustness has approximately twice as large an effect when compared to declining resilience; decreasing robustness also has a stronger and significant effect on survival. While much of the focus in previous work has been on the decline of resilience in aging (Ukraintseva et al., 2021), our results indicate that both decreasing robustness and decreasing

resilience are important processes underlying the increasing accumulation of health-related deficits with age, and the increasing rate of accumulation at older ages.

In the current study, the observed damage is thought to occur due to natural stochastic transitions, rather than a specific identified stressor (Kirkland et al., 2016; Colon-Emeric et al., 2020). Resilience measured by the observed repair also occurs without interventions (certainly in mice, due to their absence of health-care), and so is likely to represent intrinsic resilience with respect to spontaneous damage. This natural resilience can be thought of as resilience to the natural stressors of life, which continually occur during aging. While errors in deficit assessment could contribute to the damage or repair assessment, we would expect such errors to be constant with age. In contrast, we observe decreasing repair rates and increasing damage rates with age. Therefore, these age dependent rates signify decreasing resilience and robustness with age in both mice and humans.

On average, deficits accumulate more rapidly as individuals age (Mitnitski et al., 2001; Mitnitski et al., 2005; Mitnitski et al., 2013; Hoogendijk et al., 2018). Nevertheless, individual improvement as well as decline has been observed (Kojima et al., 2019; Shi et al., 2021). Previous work has modeled the change in count of deficits from baseline to a follow-up (Mitnitski et al., 2006; 2007; 2012; 2014), however that work only modeled the mean number of deficits – so that damage and repair rates were not directly assessed.

Our approach has some similarities with other recent approaches to measuring resilience or robustness. In Pyrkov et al. (2021), resilience is measured by the ability of the organism to recover from stochastic fluctuations by computing the characteristic timescale of intrinsic variations of physiological state variables, as assessed by the autocorrelation. Other studies have also used auto-correlation as a measure of resilience (Gijzel et al., 2017; Rector et al., 2021). In Arbeev et al. (2019), the onset of disease is used to indicate a lack of robustness. In these approaches, as in our approach, spontaneous deviations away from the healthy state are assessed. Conversely, some approaches measure resilience by the recovery after an acute stressor such hip

fracture or viral respiratory infection (Colon-Emeric et al., 2021). An advantage to our approach is that we observe both resilience and robustness using similar methods on the same data, so we can compare their relative effects.

One caveat with our approach is that we may miss fast damage and repair dynamics that occur on time-scales shorter than the separation between observed time-points, e.g., we cannot observe daily or weekly changes in deficit states. Therefore, our measurements of damage and repair should only be interpreted as the net damage and net repair between observed time-points. Our approach therefore results in summary measures of damage and repair rates. Nevertheless, we assess these summary measures against both age and interventions.

We find that damage and repair processes are targeted by interventions in mice. As a result, developing interventions to target either damage or repair separately is conceivable. While targeting either would affect net deficit accumulation, we found that the damage rate has a stronger effect on both mortality and the acceleration of damage accumulation than the repair rate. We predict that interventions that facilitate robustness (resistance to damage) may be more important at older ages, where damage accumulation normally accelerates. More broadly, rather than just targeting deficit accumulation or FI (Howlett et al., 2021), our results indicate that interventions could be improved by targeting an appropriate balance of damage and repair processes – in an age and sex dependent manner. Since both damage and repair occur on long timescales, this raises the possibility that these rates could be manipulated by interventions over a similarly broad range of timescales from the shortest times to organismal lifetimes. How to optimally deploy available interventions is not yet clear.

The effects of age on both damage and repair, in mice and humans, are qualitatively similar in male and female populations. Nevertheless, we have found that systemic interventions can have qualitatively distinct sex effects in mice. The ACE inhibitor enalapril has stronger effects in female mice. Voluntary exercise stopped the decline in repair rate with age for female mice, but not male mice, and stopped the increase in damage rate with age for male mice, but not female

mice. These differences suggest that assessing both damage and repair rates, together with accumulated damage as a FI, in interventional aging studies can provide a clearer assessment of sex differences. Further studies are needed to tease out the sex-dependent effects of other aging interventions, and to provide quantitative insight into the mortality-morbidity paradox, where females live longer but have higher FI scores than males (Kane and Howlett, 2021; Oksuzyan et al., 2008).

Summary measures of health such as the FI exhibit an accelerating accumulation of health deficits with age (Mitnitski et al., 2001; 2005; 2012; 2013). This universally observed behavior must be reflected in either increasing damage rates with age, or decreasing repair rates (Pyrkov et al., 2021), or – as we find – both. Both increasing damage and decreasing repair rates with age are qualitatively consistent with common theories of aging (Kirkwood et al., 2005; Gems et al., 2013; Fried et al., 2021). However, the question of whether, and by what mechanisms, damage and repair processes are coupled during aging remains unanswered. Both damage and repair rates have been typically modelled as functions of the health state in descriptive models of aging (Taneja et al., 2016; Farrell et al., 2016; Farrell et al., 2018), but without a mechanistic relationship between them apart from that imposed statistically by the observed accumulated damage. The precise relationships – and whether they are a universal feature of all aging organisms – remains to be determined. Further studies of interventions should prove useful in this regard, because they separately target damage and repair.

Our observations that repair timescales are broadly distributed, up to lifespan-scales, raise three fundamental questions for resilience studies. First, are interventions that facilitate recovery similarly effective after a broad range of timescales? This would imply that we may be able to target resilience with interventions over a longer timeframe than just acutely when damage occurs. Second, what determines the recovery timescales? As we have shown (Supplementary Figs. 5 and 7), different health attributes can have quite different recovery times. Third, would a similar broad range of resilience timescales be observed for challenge experiments with an induced stressor, and how might that depend on the magnitude and scale of the damage? It is possible that dichotomized deficits, which we have used, probe qualitatively different timescales

than continuous measures often considered in resilience studies. Future experimental resilience studies across a range of health attributes should explore longer timescales. It will be important to assess how the broad range of recovery timescales we have uncovered compare to timescales extracted from auto-correlations of physiological state variables – which have been limited to shorter times (Pyrkov et al., 2021). It is encouraging that targeted interventions may be possible long after damage has occurred – though secondary damage accumulation may nevertheless limit the benefits of such late repair.

The increasing availability of longitudinal health data over the lifespan of model aging organisms facilitates the analysis of damage and repair rates, and how they extend and change over the organismal lifespan. These damage and repair rates underlie the accumulation of damage that describes aging. Here we have shown the value of considering both resilience and robustness over the lifespan. Further studies will be able to determine how widespread organismal and sex differences in these effects are, and how universal they may prove to be. Studies of the effects on damage and repair rates of both targeted and systemic interventions will also be crucial. We have studied only three interventions so far (*e.g.* enalapril and exercise in mice, and wealth in humans). There are many other possibilities, including treatment with geroprotectors (Gonzalez-Freire et al., 2020) and lifestyle interventions, that could be deployed both in humans and in aging animal models.

References

1. Arbeev, K. G. *et al.* "Physiological Dysregulation" as a promising measure of robustness and resilience in studies of aging and a new indicator of preclinical disease. *Journals Gerontology Ser* **74**, 462–468 (2019).
2. Bisset ES, Heinze-Milne S, Grandy SA, Howlett SE. Aerobic exercise attenuates frailty in aging male and female C57Bl/6 mice and affects systemic cytokines differentially by sex. *J Gerontol A Biol Sci Med Sci*. (2021). doi: 10.1093/gerona/glab297.
3. Colón-Emeric, C. *et al.* Two approaches to classifying and quantifying physical resilience in longitudinal data. *Journals Gerontology Ser* **75**, 731–738 (2020).
4. Fried, L. P., C. M. Tangen, J. Walston, A. B. Newman, C. Hirsch, J. Gottdiener, T. Seeman, et al. "Frailty in Older Adults: Evidence for a Phenotype." *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences* 56 (3): M146–56 (2001).
5. Fried, L. P. *et al.* The physical frailty syndrome as a transition from homeostatic symphony to cacophony. *Nature Aging* **1**, 36–46 (2021).
6. Gijzel, S. M. W. *et al.* Dynamical resilience indicators in time series of self-rated health correspond to frailty levels in older adults. *Journals Gerontology Ser* **72**, 991–996 (2017).
7. Gijzel, Sanne M. W., Heather E. Whitson, Ingrid A. van de Leemput, Marten Scheffer, Dieneke van Asselt, Jerrald L. Rector, Marcel G. M. Olde Rikkert, and René J. F. Melis. "Resilience in Clinical Care: Getting a Grip on the Recovery Potential of Older Adults." *Journal of the American Geriatrics Society* 67 (12): 2650–57 (2019).
8. Gonzalez-Freire M, Diaz-Ruiz A, Hauser D, Martinez-Romero J, Ferrucci L, Bernier M, de Cabo R. The road ahead for health and lifespan interventions. *Ageing Res Rev.* (59):101037. (2020). doi: 10.1016/j.arr.2020.101037.
9. Gordon EH, Hubbard RE. Differences in frailty in older men and women. *Med J Aust.* 212(4):183-188. (2020). doi: 10.5694/mja2.50466.
10. Hadley, E. C. *et al.* Report: NIA workshop on measures of physiologic resiliencies in human aging. *The Journals of Gerontology: Series A* **72**, 980–990 (2017).
11. Hoogendijk, E. O. *et al.* Tracking changes in frailty throughout later life: results from a 17-year longitudinal study in the Netherlands. *Age and Ageing* **47**, 727–733 (2018).

12. Hoogendijk, E. O. *et al.* Frailty: implications for clinical practice and public health. *Lancet* **394**, 1365–1375 (2019).

13. Howlett, Susan E., and Kenneth Rockwood. "New Horizons in Frailty: Ageing and the Deficit-Scaling Problem." *Age and Ageing* 42 (4): 416–23. (2013).

14. Howlett, S. E., Rutenberg, A. D. & Rockwood, K. The degree of frailty as a translational measure of health in aging. *Nat Aging* **1**, 651–665 (2021).

15. Kane AE, Keller KM, Heinze-Milne S, Grandy SA, Howlett SE. 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 (2019). doi: 10.1093/gerona/gly117.

16. Kane, A. E. & Howlett, S. E. Sex differences in frailty: Comparisons between humans and preclinical models. *Mech Ageing Dev* **198**, 111546 (2021).

17. Keller K, Kane A, Heinze-Milne S, Grandy S. A, Howlett S. E. Chronic treatment with the ACE inhibitor enalapril attenuates the development of frailty and differentially modifies pro- and anti-inflammatory cytokines in aging male and female C57BL/6 mice. *J Gerontol A Biol Sci Med Sci.* **74**, 1149-1157 (2019). doi: 10.1093/gerona/gly219.

18. Kennedy BK, Berger SL, Brunet A, Campisi J, Cuervo AM, Epel ES, Franceschi C, Lithgow GJ, Morimoto RI, Pessin JE, Rando TA, Richardson A, Schadt EE, Wyss-Coray T, Sierra F. Geroscience: linking aging to chronic disease. *Cell.* **159**, 709-713 (2014). doi: 10.1016/j.cell.2014.10.039.

19. Kirkland, J. L., Stout, M. B. & Sierra, F. Resilience in aging mice. *Journals Gerontology Ser* **71**, 1407–1414 (2016).

20. Kojima, G., Taniguchi, Y., Iliffe, S., Jivraj, S. & Walters, K. Transitions between frailty states among community-dwelling older people: A systematic review and meta-analysis. *Ageing Res Rev* **50**, 81–88 (2019).

21. Kriete, A. Robustness and aging – A systems-level perspective. *Biosystems* **112**, 37–48 (2013).

22. Mitnitski A, Mogilner A, Rockwood K. Accumulation of deficits as a proxy measure of aging. *Sci. World J.* **1,** 323-336 (2001). doi:10.1100/tsw.2001.58.

23. Mitnitski, Arnold, Xiaowei Song, Ingmar Skoog, G. A. Broe, Jafna L. Cox, Eva Grunfeld, and Kenneth Rockwood. "Relative Fitness and Frailty of Elderly Men and Women in Developed Countries and Their Relationship with Mortality." *Journal of the American Geriatrics Society* 53 (12): 2184–89. (2005).

24. Mitnitski, A., Bao, L. & Rockwood, K. Going from bad to worse: A stochastic model of transitions in deficit accumulation, in relation to mortality. *Mech Ageing Dev* **127**, 490–493 (2006).

25. Mitnitski, A., Song, X. & Rockwood, K. Improvement and decline in health status from late middle age: Modeling age-related changes in deficit accumulation. *Exp Gerontol* **42**, 1109–1115 (2007).

26. Mitnitski, Arnold, Nader Fallah, Yougui Wu, Kenneth Rockwood, and Amy R. Borenstein. "Changes in Cognition during the Course of Eight Years in Elderly Japanese Americans: A Multistate Transition Model." *Annals of Epidemiology* 20 (6): 480–86. (2010).

27. Mitnitski A, Song X, Rockwood K. Trajectories of changes over twelve years in the health status of Canadians from late middle age. *Exp Gerontol.* **47**, 893-899 (2012). doi: 10.1016/j.exger.2012.06.015.

28. Mitnitski, Arnold, Xiaowei Song, and Kenneth Rockwood. "Assessing Biological Aging: The Origin of Deficit Accumulation." *Biogerontology* 14 (6): 709–17. (2013).

29. Mitnitski, A. B., Fallah, N., Dean, C. B. & Rockwood, K. A multi-state model for the analysis of changes in cognitive scores over a fixed time interval. *Stat Methods Med Res* **23**, 244–256 (2014).

30. Oksuzyan, Anna, Knud Juel, James W. Vaupel, and Kaare Christensen. "Men: Good Health and High Mortality. Sex Differences in Health and Aging." *Aging Clinical and Experimental Research* **20** (2): 91–102 (2008).

31. Phelps, A., Marmot, M., Oskala, A., Clemens, S., Banks, J., Rogers, N., Steptoe, A., Blake, M., Nazroo, J., Oldfield, Z. English Longitudinal Study of Ageing: Waves 0-9, 1998-2019. [data collection]. 31st Edition. UK Data Service. SN: 5050 (2020). doi: 10.5255/UKDA-SN-5050-24

32. Pyrkov, T. V. *et al.* Longitudinal analysis of blood markers reveals progressive loss of resilience and predicts human lifespan limit. *Nat Commun* **12**, 2765 (2021).

33. Rector, J. L. *et al.* Dynamical indicators of resilience from physiological time series in geriatric inpatients: Lessons learned. *Exp Gerontol* **149**, 111341 (2021).

34. Rockwood K, Blodgett JM, Theou O, Sun MH, Feridooni HA, Mitnitski A, Rose RA, Godin J, Gregson E, Howlett SE. A Frailty Index Based On Deficit Accumulation Quantifies Mortality Risk in Humans and in Mice. *Sci Rep.* **7**, 43068 (2017). doi: 10.1038/srep43068.

35. Scheffer, M. *et al.* Quantifying resilience of humans and other animals. *Proc National Acad Sci* **115**, 201810630 (2018).

36. Schultz, M. B. *et al.* Age and life expectancy clocks based on machine learning analysis of mouse frailty. *Nat Commun* **11**, 4618 (2020).

37. Shi, S. M., Olivieri-Mui, B., McCarthy, E. P. & Kim, D. H. Changes in a Frailty Index and Association with Mortality. *J Am Geriatr Soc* **69**, 1057–1062 (2021).

38. Sierra, F. The Emergence of Geroscience as an Interdisciplinary Approach to the Enhancement of Health Span and Life Span. *Csh Perspect Med* **6**, a025163 (2016).

39. Sierra, F. *et al.* Moving geroscience from the bench to clinical care and health policy. *J Am Geriatr Soc* (2021). doi:10.1111/jgs.17301.

40. Ukraintseva, S. *et al.* Decline in biological resilience as key manifestation of aging: Potential mechanisms and role in health and longevity. *Mech Ageing Dev* **194**, 111418 (2021).

41. Whitehead, J. C. *et al.* A Clinical Frailty Index in Aging Mice: Comparisons With Frailty Index Data in Humans. *Journals Gerontology Ser* **69**, 621–632 (2014).

42. Zhao, Xingqiu, Qiang Zhao, Jainguo Sun, and Jong S. Kim. "Generalized Log-Rank Tests for Partly Interval-Censored Failure Time Data." *Biometrical Journal.* 50 (3): 375–85 (2008).

43. Zhao, Q. "gLRT - A New R Package for Analyzing Interval-censored Survival Data", Interval-Censored Time-to-Event Data: Methods and Applications, CRC Press, 377-396 (2012).

44. Zimmer, Z., Saito, Y., Theou, O., Haviva, C. & Rockwood, K. Education, wealth, and duration of life expected in various degrees of frailty. *Eur J Ageing* **18**, 393–404 (2021).

Methods

1 Mouse data

For the mouse portion of this manuscript, published data on longitudinal health-related deficits in C57BL/6 mice from three papers was used (Keller *et al.*, 2019 [1]; Bisset *et al.*, 2021 [2]; Schultz *et al.*, 2020 [3]). A brief summary of the methods of each paper is below.

1.1 Mouse dataset 1 Keller *et al.* 2019

Male and female C57BL/6 mice were assessed for deficits approximately every 4 weeks from 16 to either 21 months of age (females) or 25 months of age (males). Mice were fed either a diet containing enalapril (280 mg/kg) or control diet for the duration of the experiment. After pre-processing (below), this data contains 21 female control mice, 25 female enalapril mice, 13 male control mice, and 25 male enalapril mice.

1.2 Mouse dataset 2 Bisset *et al.* 2021

Male and female C57BL/6 mice were assessed for deficits approximately every 2 weeks from 21 to 25 months of age. Mice were all singly housed, and half were provided a running wheel for voluntary exercise. After pre-processing (below), this data contains 11 female control mice, 11 female exercise mice, 6 male control mice, and 6 male exercise mice.

1.3 Mouse dataset 3 Schultz *et al.* 2020

Male C57BL/6Nia mice were assessed for deficits approximately every 6 weeks from 21 months of age until their natural deaths. After pre-processing (below), this data contains 44 male control mice.

1.4 Mouse clinical frailty index assessment

Each of the papers above assessed health deficits using the mouse clinical frailty index as described previously [4]. Briefly, this assessment involves scoring 31 non-invasive health-related measures in mice. Most measures are scored as 1 for a severe deficit, 0.5 for a moderate deficit and a 0 for no deficit. Deficits in body weight and temperature were scored based on deviation from reference values in young adult animals, such that a difference of less than 1 SD was scored 0, a difference of ± 1 SD was scored 0.25, a difference of ± 2 SD was scored 0.5, a difference of ± 3 SD was scored 0.75, and a difference of more than 3 SD received the maximal deficit value of 1 [4]. The deficits of malocclusions and body temperature were not assessed in mouse group 3 [3], leaving only 29 deficits for this dataset.

1.5 Mouse data pre-processing

For mouse dataset 1, we impute missing deficit values by propagating the last observed value forward. If the first observed deficit is missing, it is imputed by propagating the first observed value backward. Less than 1% of all total deficit values are missing in this dataset. No values in the other datasets are missing.

Deficits are scored on a fractional scale, with deficit i having values $d_i \in \{0, 0.25, 0.5, 0.75, 1\}$. To treat these as binary, we represent each fractional deficit d_i by 4 ordered binary deficits, $[d_i^{(1)}, d_i^{(2)}, d_i^{(3)}, d_i^{(4)}]$. Fractional deficits are then represented by setting $4 \times d_i$ of these ordered binary deficits to 1. For example if $d_i = 0.75$, this is represented as $[1, 1, 1, 0]$.

A new Frailty Index is then created by taking all of these new binary deficits, representing a $4 \times 31 = 124$ item Frailty Index ($4 \times 29 = 116$ for mouse dataset 3). This process preserves the FI scores, and a single repair or damage transition can be interpreted as taking a step of size 0.25 on the fractional deficit scale.

Measurement times with abnormally short or long intervals are removed. In mouse dataset 2, measurement times less than 0.1 months from the previous time are removed. In mouse dataset 3, measurement times more than 2 months from the previous time are removed.

In each dataset, mice with less than 2 observed time-points are removed.

2 Human data and pre-processing

We use human data from the English Longitudinal Study of Aging [5]. We select individuals that have full data for net household wealth and activities of daily living (ADL) and instrumental activities of daily living (IADL). A Frailty Index is created from the count of 10 possible ADLs and 13 possible IADLs, giving a fraction out of 23. We restrict individuals to those that were recruited to the study between the ages of 50 and 89 years. We drop individuals with follow-up time intervals above 4 years and individuals with fewer than 6 follow-ups.

We use net household wealth, as determined in the financial assessment in wave 5 of the ELSA data. We drop individuals that have parts of this assessment imputed. The raw value of net household wealth spans several orders of magnitude (and includes negatives for individuals in debt), and so is transformed by $w = \log\big(w_{\text{raw}} + \text{mean}(w_{\text{raw}})\big)$.

After pre-processing, this data contains 1049 males and 1300 females with time-intervals of approximately 2 years between observations. There are 1222 individuals from baseline ages in $[50, 60)$, 827 with baseline ages in $[60, 70)$, 281 with baseline ages in $[70, 80)$, and 19 with baseline ages in $[80, 90)$.

3 Extracting damage and repair counts

In each dataset, we observe the state of N binary health deficits $\{d_{jk}\}_{k=1}^{N}$ for each subject at a set of observation times $\{t_j\}_{j=1}^{J}$. Summing up the number of deficits at each time, we get deficit counts for each observation time, $\{n_j\}_{j=1}^{J}$, which is used to compute the Frailty Index $f_j = n_j/N$.

We compute the number of deficits damaged ($0 \to 1$ transitions) and repaired ($1 \to 0$ transitions) between two time points t_j and t_{j+1}, denoted as $n^d(t_j)$ or $n^r(t_j)$. These counts satisfy $n(t_{j+1}) = n(t_j) + n^d(t_j) - n^r(t_j)$, linking these damage and repair processes with the Frailty Index.

4 Modelling

We model deficit repair and damage as Poisson point processes with time-dependent rates, $\lambda^r(t)$ and $\lambda^d(t)$. The count of deficits repaired or damaged in an interval $[t_1, t_2]$ is assumed to follow a Poisson distribution, with mean equal to the instantaneous rate $\lambda^r(t)$ or $\lambda^d(t)$ integrated over this interval times the number of possible deficits available to be repaired $\Lambda^r(t) = \int \lambda(t)^r n_t dt$, or damaged $\Lambda^d(t) = \int \lambda^d(t)(N - n_t)dt$. We assume these time intervals are small so that we use constant rates within each time-interval to approximate these integrals, $\Lambda^r(t) \approx \lambda^r(t) n_t \Delta t$ or $\Lambda^d(t) \approx \lambda^d(t)(N - n_t)\Delta t$.

4.1 Joint longitudinal-survival model for mice data

We use a joint modelling framework to model repair and damage rates, while assessing their effect on survival. We decompose the multivariate joint distribution of the observed longitudinal damage and repair counts and survival outcome by coupling survival with the repair and damage rates $\lambda_i^r(t)$ and $\lambda_i^d(t)$ [6, 7],

$$p(T_i, c_i, \{n_i^r(t)\}, \{n_i^d(t)\} | \lambda_i^r(t), \lambda_i^d(t)) = p(T_i, c_i | \lambda_i^r(t), \lambda_i^d(t)) p(\{n_i^r(t)\}, \{n_i^d(t)\} | \lambda_i^r(t), \lambda_i^d(t)).$$

4.1.1 Longitudinal component

We use a linear Poisson model for the longitudinal damage and repair rates. A Softplus function, $\log(1 + e^x)$, is used to enforce positive rates. This function is chosen because it is approximately linear for larger values of x, in contrast to e^x which is often used for Poisson models (which resulted in poor behaviour for our models). The form of this model is,

$$\lambda_i^r(t) = \text{Softplus}\big(\boldsymbol{\beta}^r \cdot \mathbf{x}_i(t) + b_{i,0}^r + b_{i,1}^r t\big), \tag{1}$$

$$\lambda_i^d(t) = \text{Softplus}\big(\boldsymbol{\beta}^d \cdot \mathbf{x}_i(t) + b_{i,0}^d + b_{i,1}^d t\big), \tag{2}$$

$$n_i^r(t_j) | \lambda_i^r(t_j), n_i(t_j) \sim \text{Poisson}\Big(n_i(t_j)\lambda_i^r(t_j)(t_{j+1} - t_j)\Big), \tag{3}$$

$$n_i^d(t_j) | \lambda_i^d(t_j), n_i(t_j) \sim \text{Poisson}\Big((N - n_i(t_j))\lambda_i^d(t_j)(t_{j+1} - t_j)\Big), \tag{4}$$

$$n_i(t_{j+1}) = n_i(t_j) + n_i^d(t_j) - n_i^r(t_j). \tag{5}$$

The first two equations describe the time-dependent repair and damage rates, $\lambda^r(t)$ and $\lambda^d(t)$. These rates represent the probability of repair or damage, per deficit per unit time. These rates are multiplied by the number of deficits that can repair $n(t_j)$ or the number that can damage $N - n(t_j)$ and the time-interval $t_{t_{j+1}} - t_j$ to compute the mean count of repaired or damaged deficits for Poisson distributions. The last Equation 5 shows how we can compute the total count of deficits from this model, allowing this model to be used to model the Frailty Index as well.

The full-cohort parameters are denoted $\boldsymbol{\beta}$ and the subject-specific intercept and time-slopes $b_{i,0}, b_{i,1}$. The variables $\mathbf{x}_i(t)$ include the covariates and their interactions with sex and intervention group,

$$\mathbf{x}_i(t) = (1,\, t,\, \text{sex},\, \text{treatment},\, f,\, a_0,\, \text{sex} \times \text{treatment}, \text{sex} \times t,\, \text{treatment} \times t, \text{sex} \times \text{treatment} \times t).$$

The "treatment" variable is a 0/1 indicator for enalapril in mouse group 1 or exercise in mouse group 2. The other variables are the time from baseline t, the Frailty Index f, baseline age a_0, and sex (M/F). These interactions allow sex and intervention group specific time-slopes.

The repair and damage processes are linked by including correlations between the subject-specific parameters $[\boldsymbol{b}_i^r, \boldsymbol{b}_i^d] \sim \mathcal{N}(0, \Sigma)$.

4.1.2 Survival component

We jointly model these repair and damage processes with survival, with proportional hazards and a baseline hazard parameterized with M-splines [8] (which are always non-negative). The damage and repair processes are linked with survival by including damage rate and repair rate in the hazard rate,

$$\begin{aligned} h_i(t) &= h_0(t, \text{sex}) \exp\Big(\boldsymbol{\gamma} \cdot \mathbf{u}_i(t) + \gamma^r \text{Softplus}^{-1} \lambda_i^r(t) + \gamma^d \text{Softplus}^{-1} \lambda_i^d(t)\Big), && (6) \\ h_0(t) &= (\text{male}) \cdot \sum_{l=1}^{L} a_{l,\text{male}} M_{l,3}(t|\mathbf{k}) + (\text{female}) \cdot \sum_{l=1}^{L} a_{l,\text{female}} M_{l,3}(t|\mathbf{k}), \;\; \sum_{l=1}^{L} a_l = 1, \;\; a_l \geq 0, \\ S_i(t) &= \exp\big(-\int_{t_0}^{t} h_i(s) ds\big). && (7) \end{aligned}$$

The first equation describes the hazard rate $h_i(t)$ in terms of the covariates $\mathbf{u}_i$ and the repair and damage rates. The baseline hazard $h_0(t, \text{sex})$ is modeled with sex-specific splines in Equation 7, due to the large disparity in survival by sex. The covariates are $\mathbf{u}_i = (1,\, \text{sex},\, \text{treatment},\, \text{sex} \times \text{treatment},\, f,\, a_0)$.

4.1.3 Priors and hyperparameters

We use weakly informative priors to regularize parameters,

$$\beta_0^r,\; \beta_0^d,\; \gamma_0 \sim \mathcal{N}(0, 3), \quad \boldsymbol{\beta}^r,\; \boldsymbol{\beta}^d,\; \boldsymbol{\gamma},\; \gamma^r,\; \gamma^d \sim \mathcal{N}(0, 1), \tag{8}$$

$$[\mathbf{b}_i^r,\; \mathbf{b}_i^d] \sim \mathcal{N}(0, \boldsymbol{\Sigma}), \quad \boldsymbol{\Sigma} = \boldsymbol{\sigma}\boldsymbol{\Omega}\boldsymbol{\sigma}, \quad \boldsymbol{\sigma} \sim \text{HalfCauchy}(0, 1), \quad \boldsymbol{\Omega} \sim \text{LKJ}(2), \tag{9}$$

$$\mathbf{a} \sim \text{Dirichlet}(1.0, L = 17), \quad \boldsymbol{\kappa} = \{\min(\{T_i\}_i), Q_{0.05}(\{T_i\}_i), ..., Q_{0.95}(\{T_i\}_i), \max(\{T_i\}_i)\}. \tag{10}$$

Broad $\mathcal{N}(0, 3)$ priors are used on intercept parameters and narrow $\mathcal{N}(0, 1)$ priors on covariate coefficients (Equation 8). The covariance matrix $\boldsymbol{\Sigma}$ for the coupling of the subject-specific parameters $\mathbf{b}_i$ is decomposed in terms of a correlation matrix $\boldsymbol{\Omega}$ with a LKJ prior and standard deviations $\boldsymbol{\sigma}$ with half-Cauchy distributions (Equation 9). Spline coefficients $\mathbf{a}$ use a Dirichlet distribution with concentration 1, representing a uniform prior on the simplex $\sum_{l=1}^{L} a_l = 1,\; a_l \geq 0$. We use $L = 17$ spline knots with knots at the minimum last follow-up age, the maximum, and 15 uniformly spaced quantiles from 0.05 to 0.95 of the last follow-up age (Equation 10).

Integrals of the hazard rate are computed with 5-point Gaussian Quadrature between each observed time interval.

4.2 Non-linear modeling for human data

There is much more human data than mice and the data is more complex, where linear effects are not sufficient to capture the combined influence of wealth, baseline age, and time. We use a non-linear Poisson model with non-constant coefficients to include addition degrees of freedom. We parameterize these non-constant coefficients with B-splines. The individuals selected from ELSA with wealth data do not have mortality data available, simplifying the model from the joint model used above for mice.

Our model has the form,

$$\lambda_i^r(t) = \text{Softplus}\big(\boldsymbol{\beta}_0^r \cdot \mathbf{x}_i(t) + \beta_1^r(w, a_0) + \beta_2^r(w, a_0) \times \text{sex} + \beta_3^r(w, a_0) \times t + \beta_4^r(w, a_0) \times \text{sex} \times t + b_{i,0}^r\big), \tag{11}$$

$$\lambda_i^d(t) = \text{Softplus}\big(\boldsymbol{\beta}_0^d \cdot \mathbf{x}_i(t) + \beta_1^d(w, a_0) + \beta_2^d(w, a_0) \times \text{sex} + \beta_3^d(w, a_0) \times t + \beta_4^d(w, a_0) \times \text{sex} \times t + b_{i,0}^d\big), \tag{12}$$

$$n_i^r(t_j) | \lambda_i^r(t_j), n_i(t_j) \sim \text{Poisson}\Big(n_i(t_j) \lambda_i^r(t_j)(t_{j+1} - t_j)\Big), \tag{13}$$

$$n_i^d(t_j) | \lambda_i^d(t_j), n_i(t_j) \sim \text{Poisson}\Big(N - n_i(t_j) \lambda_i^d(t_j)(t_{j+1} - t_j)\Big), \tag{14}$$

where

$$\mathbf{x}_i(t) = (1,\, t,\, \text{sex},\, w,\, f,\, a_0,\, \text{sex} \times t,\, w \times t,\, a_0 \times t,\, \text{sex} \times f,\, w \times f,\, a_0 \times f). \tag{15}$$

The non-constant coefficients $\{\beta_k(w, a_0)\}_k$ are implemented as 2D B-splines for wealth and baseline age with 5 wealth knots and 5 baseline age knots at the minimum, maximum and terciles of these variables. We use smoothing 2D random-walk priors on the spline coefficients,

$$s_{11}, \tau_w, \tau_{b_0} \sim \mathcal{N}(0, 1), \quad p_w, p_{b_0} \sim \text{Dirichlet}(1.5), \tag{16}$$

$$s_{ij} \sim p_w \mathcal{N}(s_{i-1,j}, \tau_w) + p_{b_0} \mathcal{N}(s_{i,j-1}, \tau_{b_0}), \tag{17}$$

$$\beta_k(w, a_0) = \sum_{i,j=1}^{5} s_{ij} B_{i,3}(w; \boldsymbol{\kappa}_w) B_{j,3}(a_0; \boldsymbol{\kappa}_{a_0}). \tag{18}$$

All other priors are the same as in the mouse modelling.

Note, in the human data we do not include subject-specific time-slopes $b^r_{i,1}$ and $b^d_{i,1}$ as we did in the mouse data, since we have much shorter time-series. When these slopes are included, we see evidence of the model over-fitting to the data by the proportion of residuals including zero within 95% credible intervals being much higher than 0.95 – nearing 0.99 to 1.00.

4.2.1 Derivatives

We can compute the derivative of the Frailty Index according to the modelled repair and damage rates,

$$\frac{d}{dt} f_i(t) = (1 - f_i)\lambda^d_i(t) - f_i \lambda^r_i(t). \tag{19}$$

To understand the effect of interventions, we compute the derivative with respect to time for the repair and damage rates,

$$\begin{aligned} \frac{d}{dt}\lambda^r_i(t) &= \frac{\partial \lambda^r_i(t)}{\partial t} + \frac{\partial \lambda^r_i(t)}{\partial f}\frac{df_i(t)}{dt}, \\ &= \Big(\boldsymbol{\beta}^r \cdot \frac{d\mathbf{x}_i(t)}{dt} + \boldsymbol{\beta}^r \cdot \frac{d\mathbf{x}_i(t)}{df} + \mathbf{b}^r_i \cdot \frac{d\mathbf{z}_i(t)}{dt}\Big)\frac{e^{\lambda^r_i(t)}}{e^{\lambda^r_i(t)} + 1}, \end{aligned} \tag{20}$$

$$\begin{aligned} \frac{d}{dt}\lambda^d_i(t) &= \frac{\partial \lambda^r_i(t)}{\partial t} + \frac{\partial \lambda^r_i(t)}{\partial f}\frac{df_i(t)}{dt}, \\ &= \Big(\boldsymbol{\beta}^d \cdot \frac{d\mathbf{x}_i(t)}{dt} + \boldsymbol{\beta}^d \cdot \frac{d\mathbf{x}_i(t)}{df} + \mathbf{b}^d_i \cdot \frac{d\mathbf{z}_i(t)}{dt}\Big)\frac{e^{\lambda^d_i(t)}}{e^{\lambda^d_i(t)} + 1}. \end{aligned} \tag{21}$$

This is the slope of these rates vs time, with the increase in the Frailty Index $f(t)$ included. While we only include explicit linear effects of time in the model, the increase in Frailty Index with time can influence the derivative to change.

We can compute the curvature as the second derivative of the Frailty Index with age, written in terms of first derivatives of the rates,

$$\frac{d^2}{dt^2} f_i(t) = \left[(1 - f_i(t))\frac{d\lambda^d_i(t)}{dt} - \frac{df_i(t)}{dt}\lambda^d_i(t)\right] - \left[f_i(t)\frac{d\lambda^r_i(t)}{dt} + \frac{df_i(t)}{dt}\lambda^r_i(t)\right]. \tag{22}$$

The first group of terms are those involving damage rate (robustness) and the second group of terms are those involving repair (resilience). These terms are plotted in Figure 3.

4.3 Repair and damage timescale mice and human data

We observe the amount of time that has passed between damage and repair events, and vice versa. This can be used to determine the time-scales of these damage and repair processes. However, since a deficit might damage and the individual dies before the deficit is ever repaired, there is right censoring. Additionally, observations are only made at specific time-points and so we cannot determine the exact time at which a deficit damaged or repaired, there is interval censoring. To estimate the distribution of repair and damage times, we treat repair and damage events for each deficit as a mixture of interval and right censored events. Accordingly, we model state-survival curves for the damaged state (time-scale of resilience) and undamaged state (time-scale of robustness).

We use a Bayesian survival model with M-splines for the baseline hazard,

$$h(t) = e^{\gamma_0} \sum_{l=1}^{L} a_l M_{l,3}(t|\mathbf{k}), \qquad \sum_{l=1}^{L} a_l = 1, \qquad S(t) = \exp\big(-\int_{t_0}^{t} h(s)ds\big). \tag{23}$$

This is fit separately for sex and control/intervention groups.

We include both interval-censoring and right-censoring in the likelihood,

$$p(T_i^L, T_i^U, T_i, c_i | \{a_l)\}_l, \gamma_0) = [S(T_i)]^{c_i} [S(T_i^L) - S(T_i^U)]^{1-c_i}, \tag{24}$$

where T^L is the lower interval bound, T^U is the upper interval bound, and T is a time of right censoring. We use 32 knots set at 30 evenly spaced quantiles of the event times from 0.1 to 0.9 together with the minimum and maximum event time. A uniform Dirichlet$(1.0, 32)$ prior is used for the spline coefficients and a broad $\mathcal{N}(0, 10)$ prior for γ_0.

4.4 MCMC Sampling

We use the STAN no U-turn sampler (NUTS) [9]. We use 4000 warm-up iterations and 6000 sampling iterations on 4 separate chains for the mouse joint models. For mouse dataset 2 we use the sampler settings adapt_delta=0.95, max_treedepth=20 to avoid divergent transitions. For the human model we use 2 separate chains with 1000 warm-up iterations and 3000 sampling iterations. For the interval-censored Bayesian survival models we use 2000 warm-up iterations and 3000 sampling iterations for 4 separate chains.

In the Supplemental Figure 1 we perform posterior predictive checks [10, 11] for the mice and human models by plotting observed and simulated distributions of counts. We also compute R^2 statistics [12] and the coverage of credible intervals for residuals.

5 Code and data availability

Our code for data pre-processing, modelling, and plotting is available `https://github.com/Spencerfar/aging-damagerepair`. Human ELSA data can be accessed by agreeing to an End User Licence `https://www.elsa-project.ac.uk/accessing-elsa-data` and downloading waves 1 to 9. Mouse dataset 3 is freely available at `https://github.com/SinclairLab/frailty` [3].

References

[1] Kaitlyn Keller, Alice Kane, Stefan Heinze-Milne, Scott A Grandy, and Susan E Howlett. Chronic treatment with the ACE inhibitor enalapril attenuates the development of frailty and differentially modifies pro- and anti-inflammatory cytokines in aging male and female C57BL/6 mice. *J. Gerontol. A Biol. Sci. Med. Sci.*, 74(8):1149–1157, July 2019.

[2] Elise S Bisset, Stefan Heinze-Milne, Scott A Grandy, and Susan E Howlett. Aerobic exercise attenuates frailty in aging male and female C57Bl/6 mice and effects systemic cytokines differentially by sex. *J. Gerontol. A Biol. Sci. Med. Sci.*, October 2021.

[3] Michael B Schultz, Alice E Kane, Sarah J Mitchell, Michael R MacArthur, Elisa Warner, David S Vogel, James R Mitchell, Susan E Howlett, Michael S Bonkowski, and David A Sinclair. Age and life expectancy clocks based on machine learning analysis of mouse frailty. *Nat. Commun.*, 11(1):1–12, September 2020.

[4] Jocelyne C Whitehead, Barbara A Hildebrand, Michael Sun, Michael R Rockwood, Robert A Rose, Kenneth Rockwood, and Susan E Howlett. A clinical frailty index in aging mice: comparisons with frailty index data in humans. *J. Gerontol. A Biol. Sci. Med. Sci.*, 69(6):621–632, June 2014.

[5] A. Phelps, M. Marmot, A. Oskala, S. Clemens, J. Banks, N. Rogers, A. Steptoe, M. Blake, J. Nazroo, and Z. Oldfield. English longitudinal study of ageing: Waves 0-9, 1998-2019. [data collection] 31st edition. *UK Data Service*, 2020.

[6] Graeme L. Hickey, Pete Philipson, Andrea Jorgenson, and Kolamunnage-Dona. Joint modelling of time-to-event and multivariate longitudinal outcomes: recent developments and issues. *BMC Medical Research Methodology*, 16(117), 2016.

[7] Samuel L. Brilleman, Eren M. Elci, Jacqueline Buros Novik, and Rory Wolfe. Bayesian survival analysis using the rstanarm R package. *arXiv*, 2002.09633, 2020.

[8] J. O. Ramsay. Monotone regression splines in action. *Statist. Sci.*, 3(4):425–441, 11 1988.

[9] Stan Development Team. The Stan core library version 2.25.0, 2020.

[10] Jonah Gabry, Daniel Simpson, Aki Vehtari, Michael Betancourt, and Andrew Gelman. Visualization in bayesian workflow. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 182(2):389–402, Jan 2019.

[11] Andrew Gelman, Aki Vehtari, Daniel Simpson, Charles C. Margossian, Bob Carpenter, Yuling Yao, Lauren Kennedy, Jonah Gabry, Paul-Christian Bürkner, and Martin Modrák. Bayesian workflow, 2020.

[12] Aki Vehtari, Andrew Gelman, Daniel Simpson, Bob Carpenter, and Paul-Christian Bürkner. Rank-normalization, folding, and localization: An improved $\widehat{R}$ for assessing convergence of MCMC. *Bayesian Analysis*, Jul 2020.

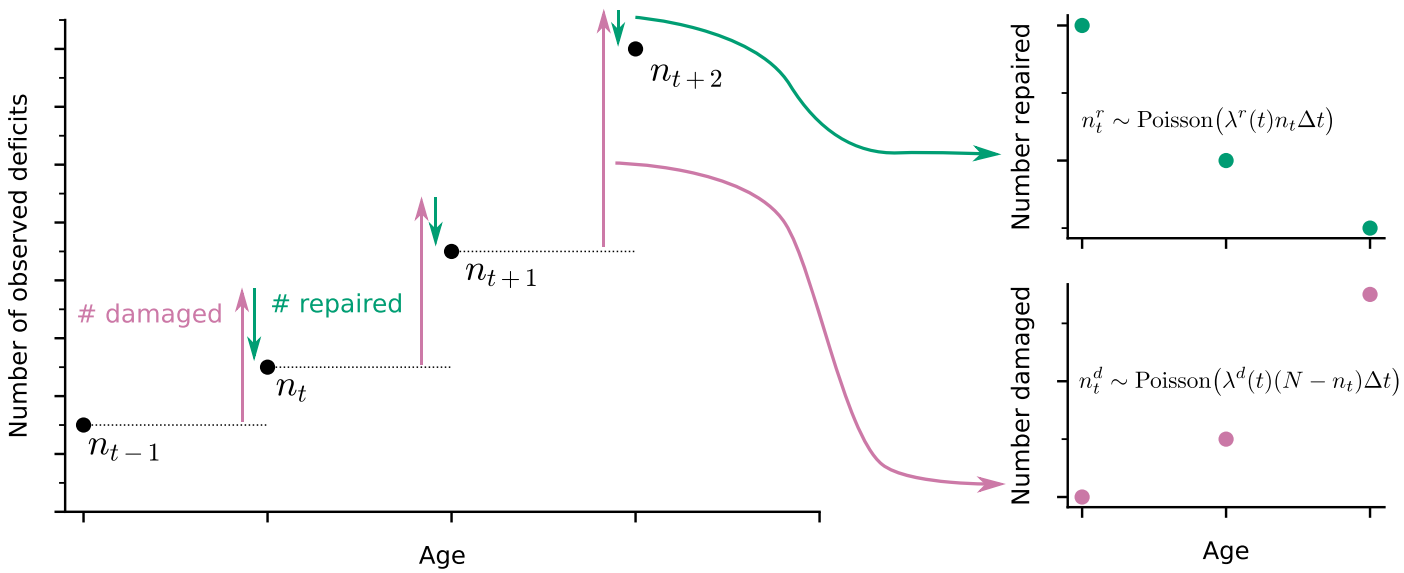


Figure 1. **Extracting damage and repair from the longitudinal observation of binary health deficits.** Instead of just considering the Frailty Index (FI) or net count of deficits at each age n_t (i.e. (FI) multiplied by the total number of deficits considered N) as a measure of health, we separately consider the number of deficits damaged n_t^d or repaired n_t^r between time intervals Δt. Time-dependent damage $\lambda^d(t)$ and repair rates $\lambda^r(t)$ are extracted with Poisson models for the counts of repaired or damaged deficits.

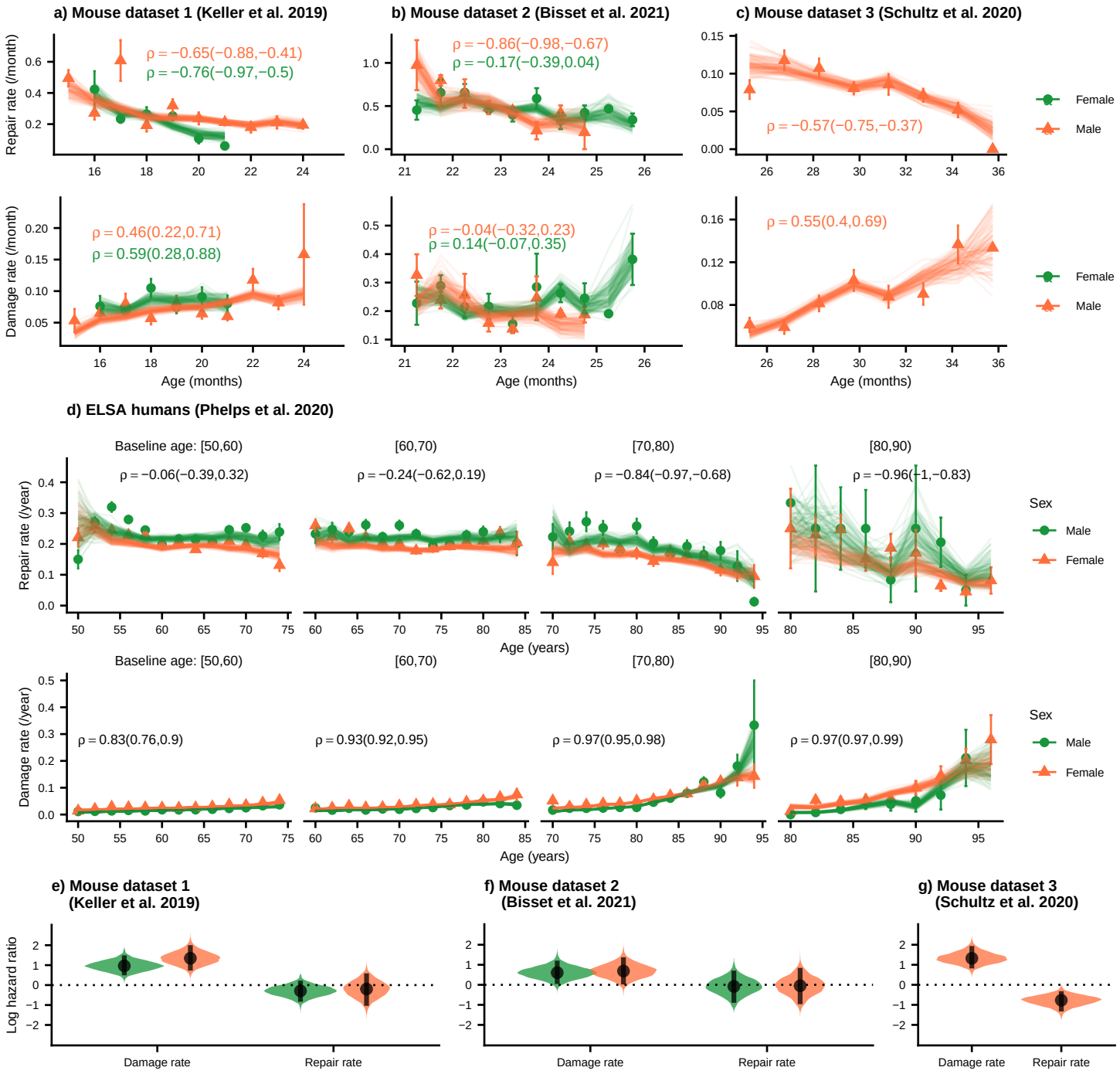


Figure 2. **Repair rates decrease and damage rates increase with age.** Repair rates vs age (top), damage rates vs age (bottom) for **a)** Mouse dataset 1 (Keller *et al.*, 2019), **b)** Mouse dataset 2 (Bisset *et al.*, 2021), **c)** Mouse dataset 3 (Schultz *et al.*, 2020) and **d)** ELSA humans (Phelps *et al.* 2020) plotted by decades of baseline age. Points in all plots represent binned averages of rates from the data with standard errors, and lines represent posterior samples from Bayesian models of the rates (see Methods). For each plot, the mean Spearman's rank correlation ρ between the rate and age is indicated by the median of the posterior and a 95% posterior credible interval in parenthesis. **e)-f)** Posterior distributions of log hazard ratios of death for damage and repair rates are shown as violin plots for the mouse datasets. These hazard ratios correspond to a 1 standard deviation increase in the damage or repair rates. The black interval shows a 95% credible interval around the median point.

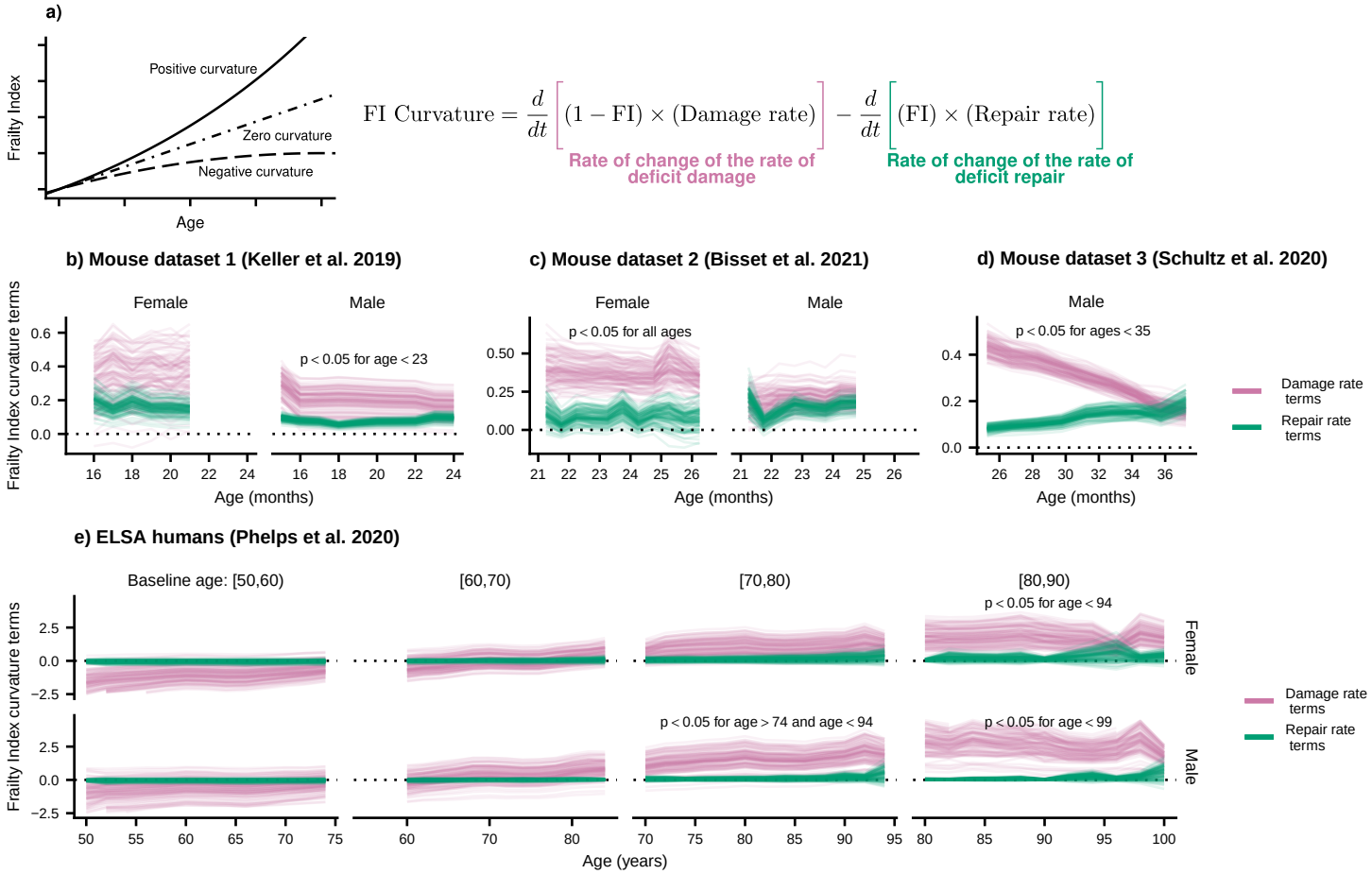


Figure 3. **Frailty Index curvature is dominated by declining robustness. a)** Frailty Index curvature measures the rate of accumulation of damage. Positive curvature indicates an acceleration of damage accumulation, zero curvature indicates a constant accumulation of damage, and negative curvature indicates a decelerating accumulation of damage. Curvature is computed with the second time-derivative of the Frailty Index (Methods Equation 31). Terms of the curvature involving the repair rate (green) and the damage rate (pink) are separately shown. Lines represent posterior samples from our Bayesian models for **b)** Mouse dataset 1 (Keller *et al.*, 2019), **c)** Mouse dataset 2 (Bisset *et al.*, 2021), **d)** Mouse dataset 3 (Schultz *et al.*, 2020) and **e)** ELSA humans (Phelps *et al.*, 2020), plotted separately by decades of baseline age. On all plots, we indicate for which ages the proportion of the posterior for the difference in these terms that is negative is below 0.05; the Bayesian analogue of a p-value testing the contributions of robustness and resilience. In Supplemental Figure 3, this can be seen with posterior credible intervals for these differences.

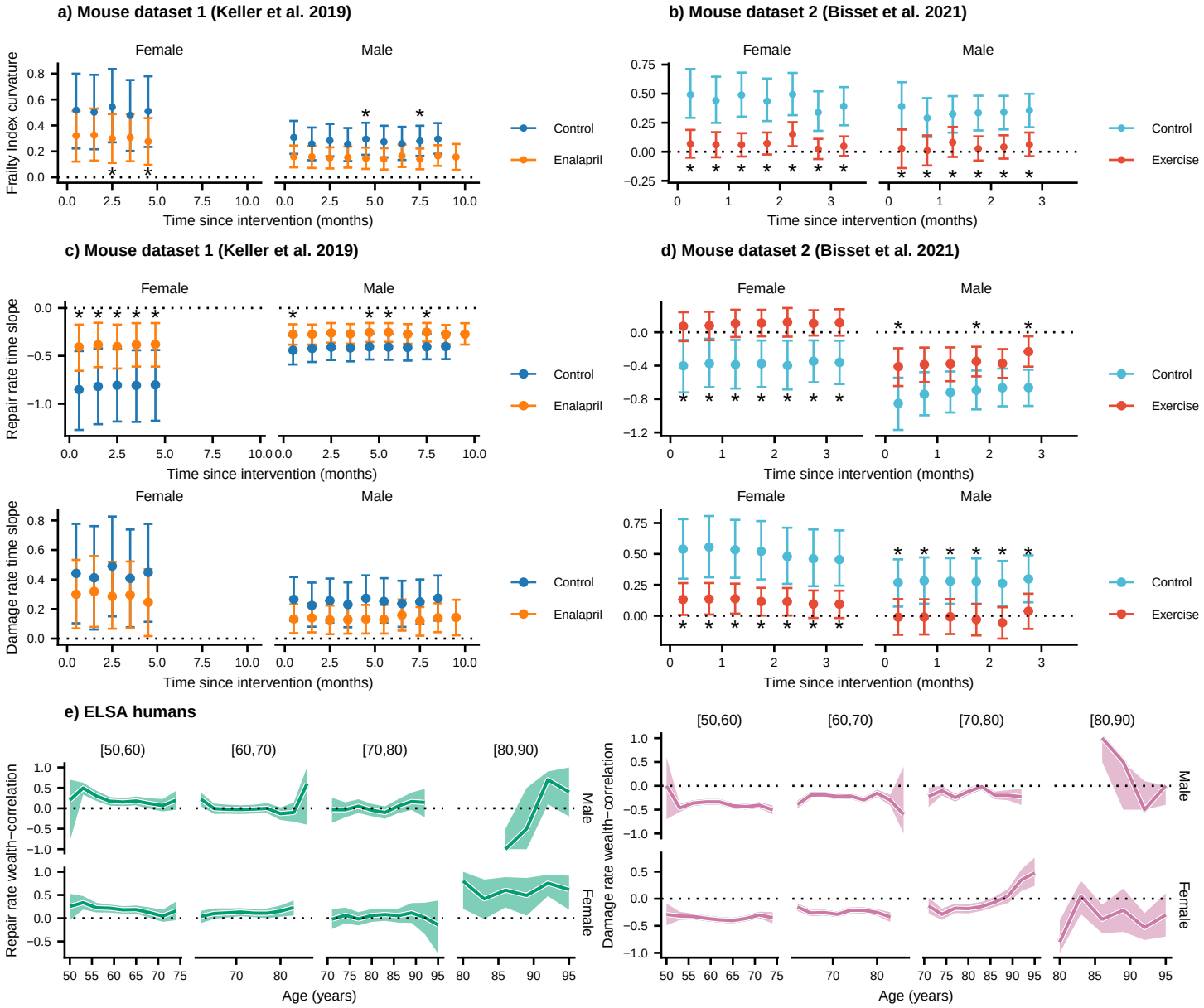


Figure 4. **Interventions both increase resilience and decrease damage. a),b)** The effect of enalapril and exercise on Frailty Index curvature for mice for mouse datasets 1 and 2. 95% credible intervals for these curvatures are shown by errorbars around the median (point). Asterisks (*) indicate credible intervals for the difference between intervention and control fully exclude zero. **c),d)** Repair rates and damage rates time-slopes vs time since intervention for the effect of enalapril and exercise for mouse datasets 1 and 2. 95% credible intervals for these curvatures are shown by errorbars around the median (point). Asterisks (*) indicate credible intervals for the difference between intervention and control fully exclude zero. **e)** Spearman rank correlation ρ between wealth and repair rate (green) and damage rate (pink), vs age for ELSA humans. Individuals are separated by decades of baseline age, and 95% credible intervals for these correlations are shown as coloured regions around the median (thick line). We restrict this plot to ages with at least 3 individuals.

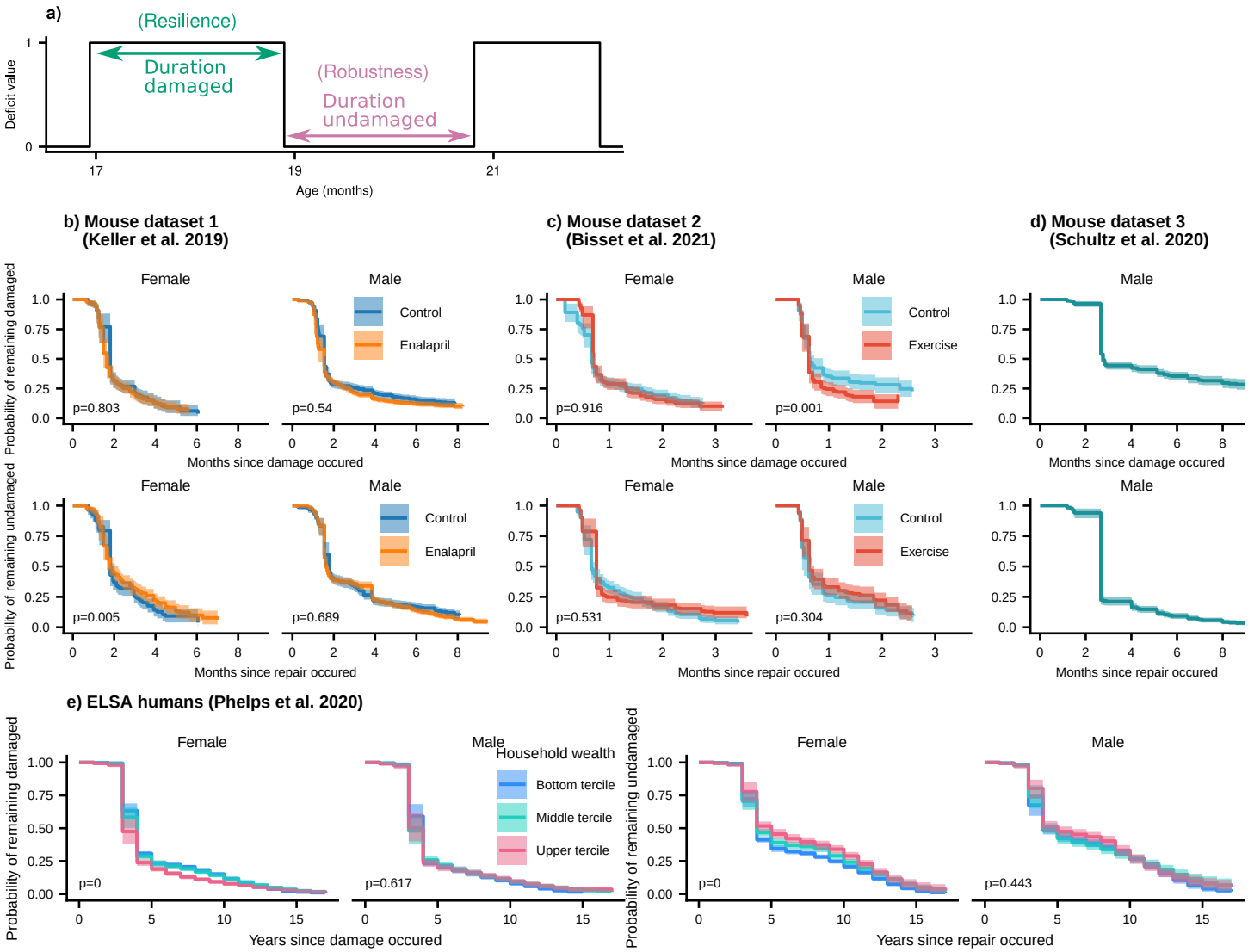
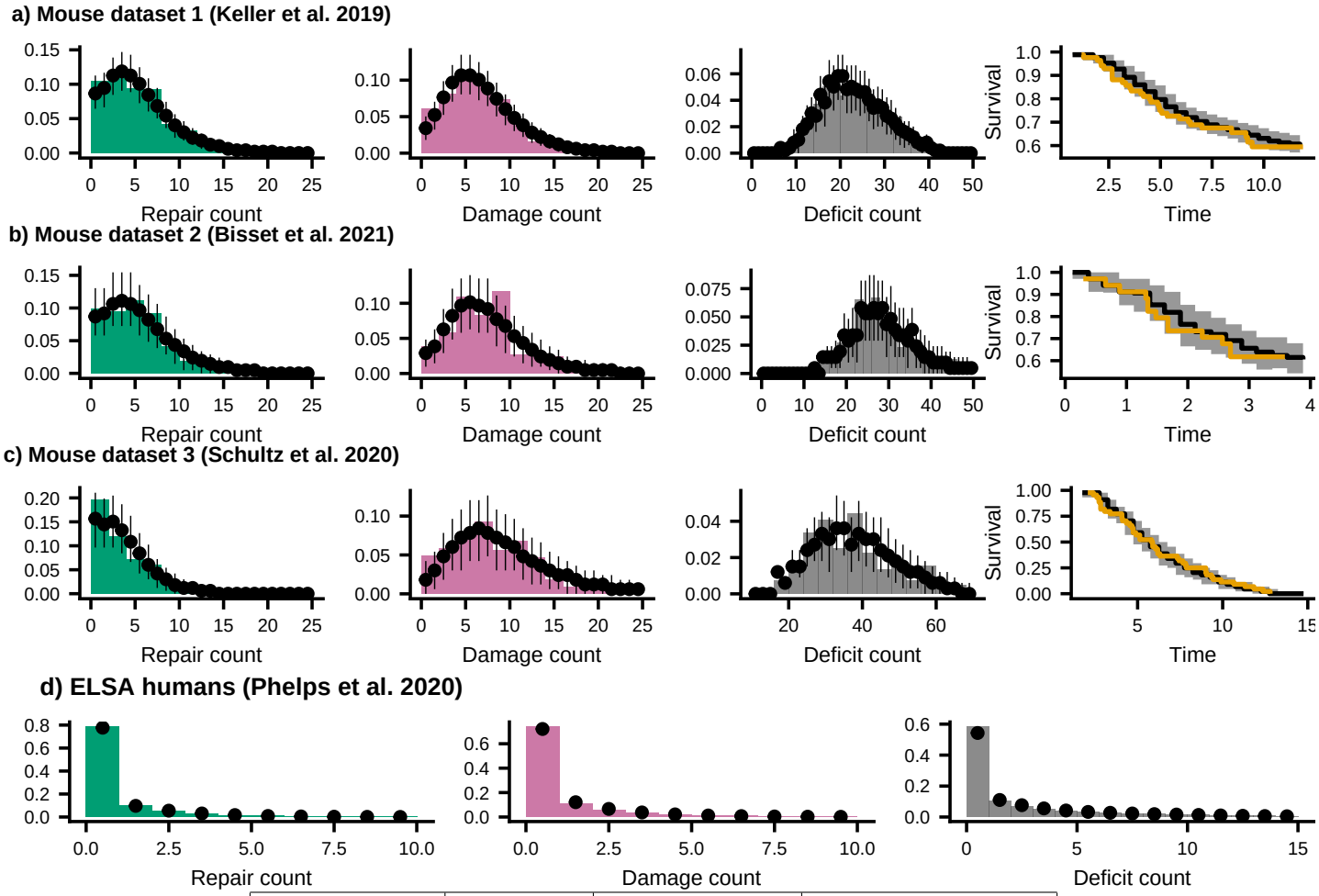


Figure 5. **Resilience and robustness occur over both short and long time-scales in both mice and humans. a)** The time-scale of resilience is measured as the lifetime of the damaged state. The time-scale of robustness is measured as the lifetime of the undamaged state. Deficit state-survival curves, showing the probability of remaining in the current damaged or undamaged state since time of transition, are shown for **c)** Mouse dataset 1, **d)** Mouse dataset 2, **e)** Mouse dataset 3, and **f)** ELSA human dataset. P-values are shown on the lower left of each plot for generalized log-rank tests for the equality of the survival functions between the intervention or wealth groups (Zhao *et al.* 2008 and 2012).



Dataset	Outcome	Median R^2 (95% CrI)	Proportion within 95% CrI
Mouse dataset 1 (Keller *et al.* 2019)	Repair counts	0.42(0.48, 0.57)	0.96
	Damage counts	0.46(0.42, 0.50)	0.91
	Deficit counts	0.48(0.45, 0.50)	0.96
Mouse dataset 2 (Bisset *et al.* 2021)	Repair counts	0.51(0.45, 0.56)	0.96
	Damage counts	0.54(0.48, 0.60)	0.94
	Deficit counts	0.50(0.46, 0.53)	0.97
Mouse dataset 3 (Schultz *et al.* 2019)	Repair counts	0.49(0.41, 0.58)	0.96
	Damage counts	0.46(0.39, 0.52)	0.95
	Deficit counts	0.48(0.45, 0.51)	0.94
ELSA humans	Repair counts	0.52(0.50, 0.53)	0.98
	Damage counts	0.44(0.43, 0.45)	0.96
	Deficit counts	0.47(0.46, 0.47)	0.95

Figure 2-figure supplement 1. **Posterior predictive check for joint models.** We compare histograms of the observed repair counts (green), damage counts (red), total deficit counts (grey), and survival probability (orange) with posterior samples from the model (black points, showing medians and 95% credible intervals) for **a)** Mouse dataset 1 (Keller *et al.*, 2019), **b)** Mouse dataset 2 (Bisset *et al.*, 2021), **c)** Mouse dataset 3 (Schultz *et al.*, 2020), and **d)** ELSA humans (Phelps *et al.* 2020). These distributions show the distributions of these counts for all individuals, all time-points, and all sex and treatment groups. The table shows R^2 statistics and the proportion of posterior 95% credible intervals (CrI) for data-points where the residual includes zero (expected to be 0.95). These tests show that the models accurately represent the observed data.

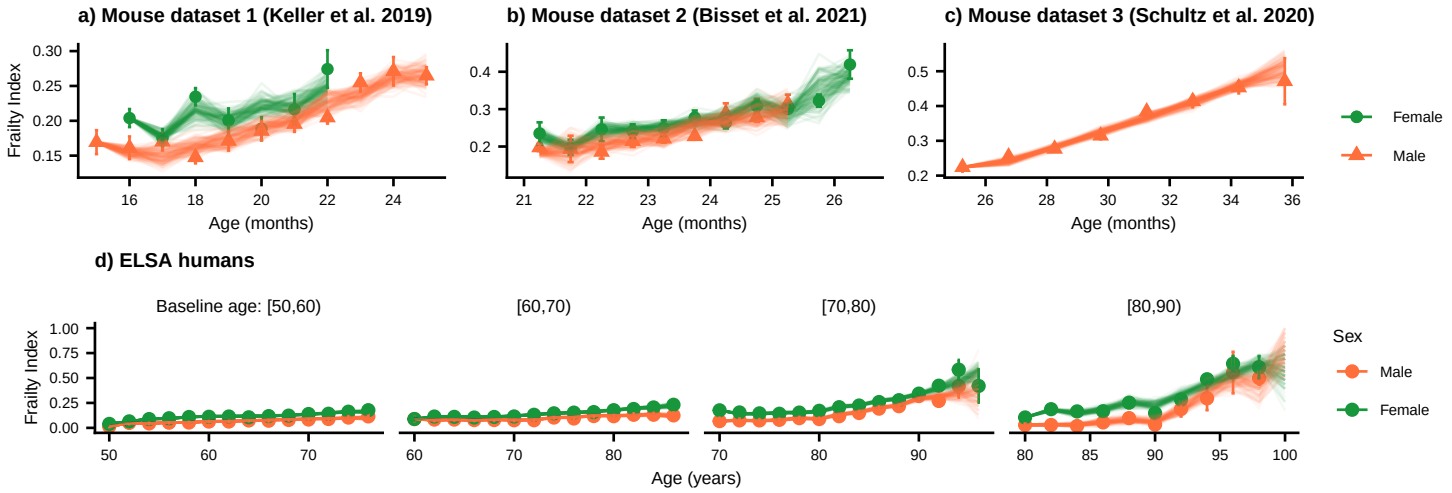


Figure 2-figure supplement 2. **Increase in Frailty Index in mice and humans.** The increase in the Frailty Index with age for **a)** Mouse dataset 1, **b)** Mouse dataset 2, **c)** Mouse dataset 3, and **d)** ELSA human dataset. We observe higher FI for females (green) than males (orange) for both mice and humans.

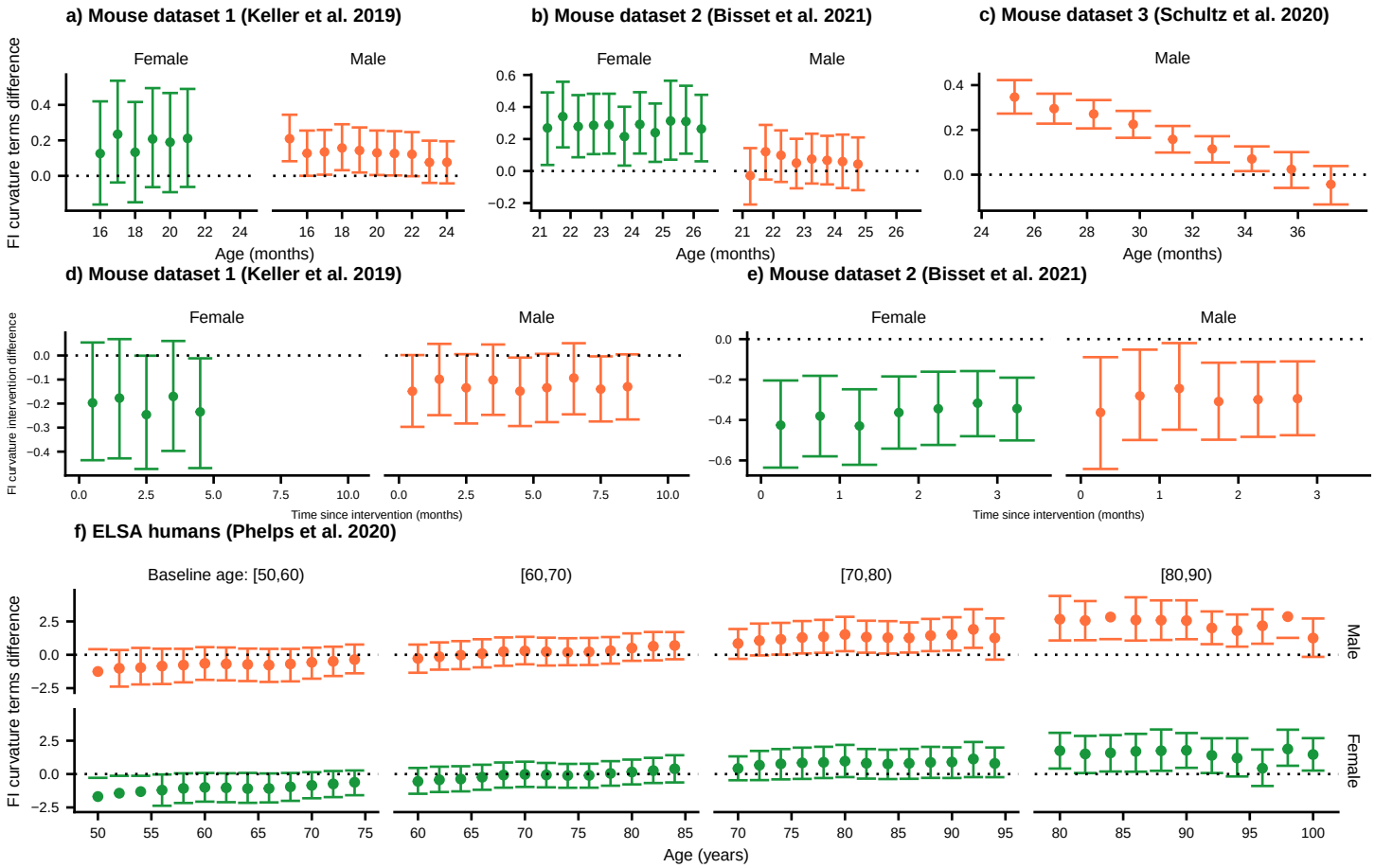


Figure 3-figure supplement 1. **Testing the effect of robustness, resilience, and interventions on curvature.** We compute the posterior distribution mean difference between the curvature terms involving damage rates and repair rates for the control. Here we show the median of the posterior (points) with 95% credible intervals for the control groups for mouse datasets **a)** 1, **b)** 2, and **c)** 3. For credible intervals above zero, the effect of the damage rate on the curvature is considered significant at the 95% level. In **d)** and **e)**, we show test the effect of enalapril and exercise on the curvature, showing that exercise strongly reduces curvature. **f)** We compute the posterior distribution mean difference between the curvature terms involving damage rates and repair rates for humans.

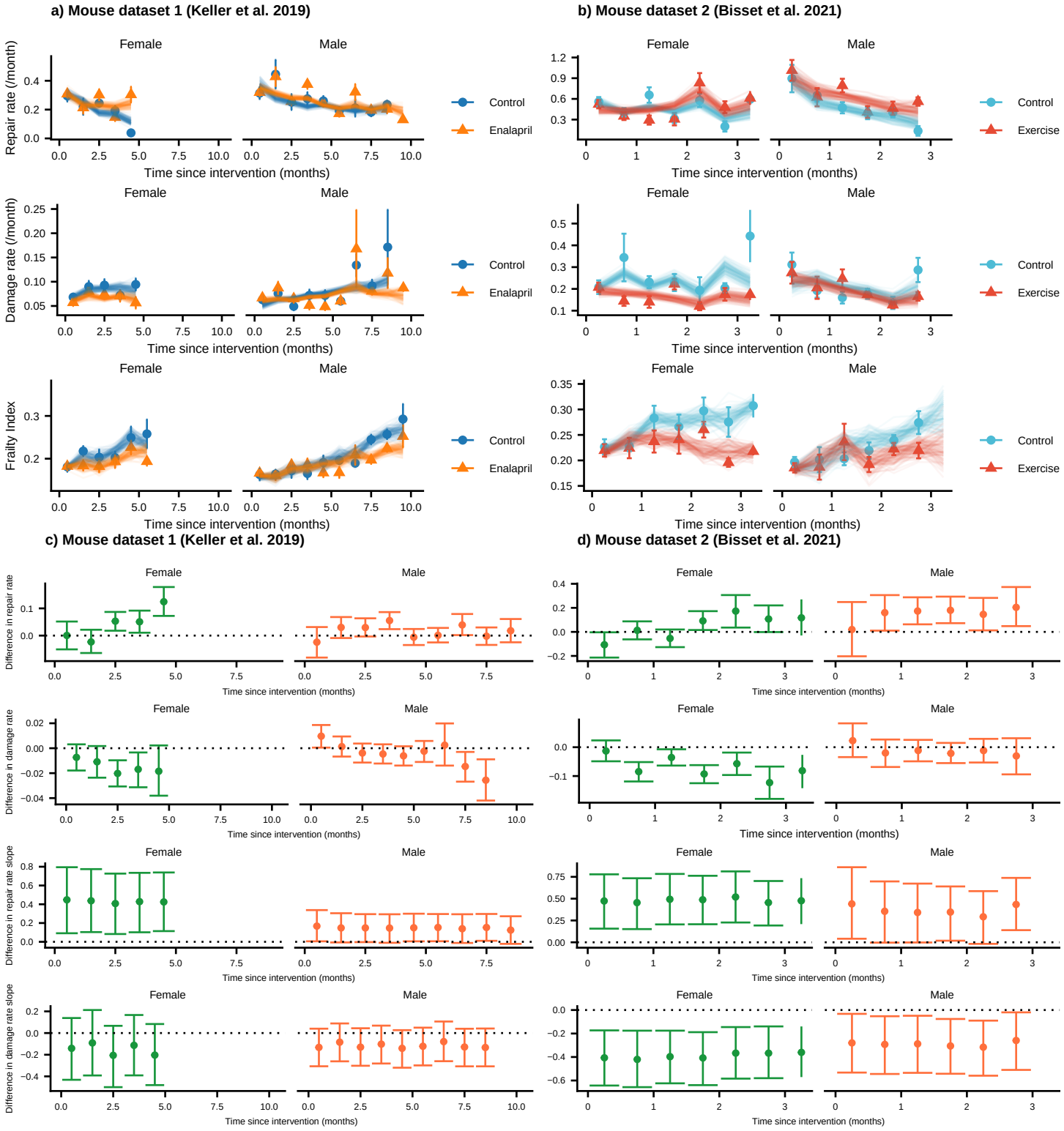


Figure 4-figure supplement 1. **The effect of interventions on repair, damage, and Frailty Index in mice. a)** Repair rates, damage rates, and Frailty Index vs age for control mice and mice treated with enalapril from Keller et al. 2019. **b)** Repair rates, damage rates, and Frailty Index vs age for control mice and mice with voluntary exercise from Bisset et al. 2021. Both enalapril and exercise impact repair and damage rates, leading to a lower Frailty Index. These plots show binned averages of the rates of Frailty Index as the points with standard errors, overlayed with posterior samples from the models. In **c)** and **d)**, we test the difference between intervention and control groups for repair rates, damage rates, repair rate time slopes, and damage rate time slopes by showing the posterior median and 95% credible intervals. Intervals that don't cross zero are considered significant at the 95% level.

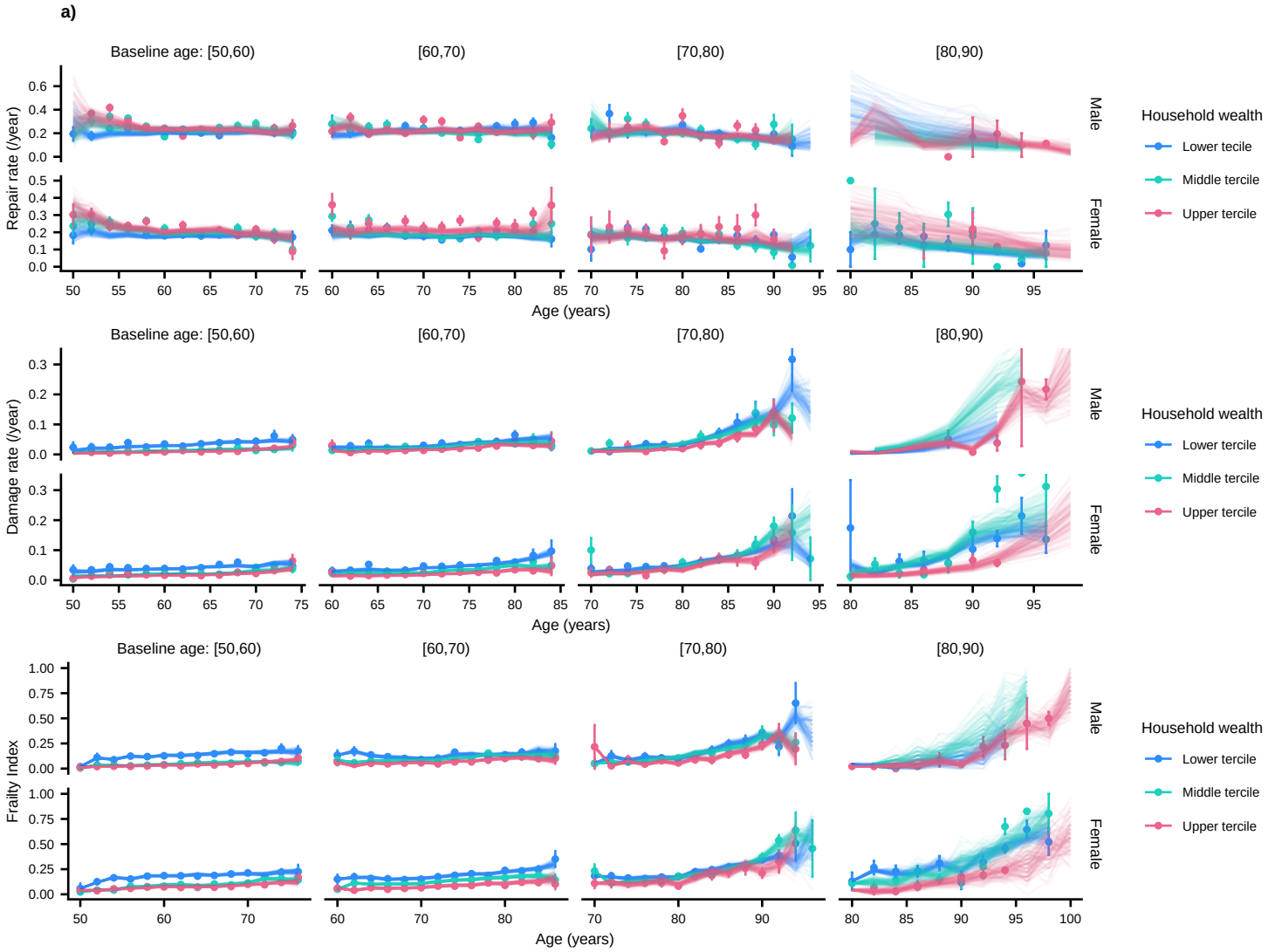


Figure 4-figure supplement 2. **Humans stratified by terciles of household wealth.** Repair rate, damage rate and Frailty Index vs age for ELSA humans (Phelps *et al.* 2020) are shown, stratified by terciles of net household wealth. These plots show binned averages of the rates of Frailty Index as the points with standard errors, overlayed with posterior samples from the models. The lowest tercile exhibits lower repair rates and higher damage rates for younger ages.

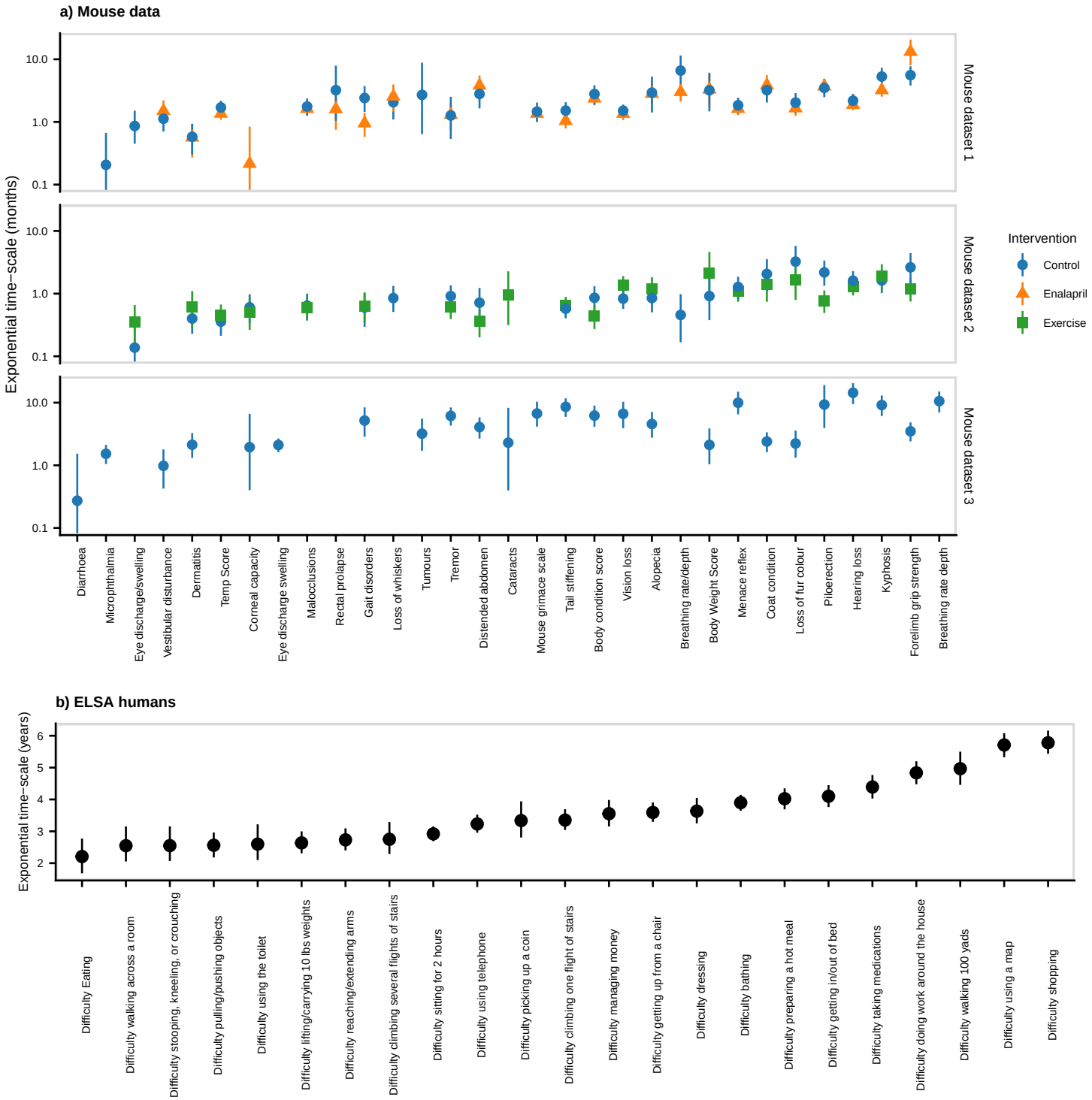


Figure 5-figure supplement 1. **Time-scales of resilience for individual deficits.** Exponential time-scales for damaged state survival curves. Exponential survival models are fit for the survival of the damaged state for each binary health variable. Time-scales are computed as the inverse of the exponential rate, and the posterior median is shown as a point, with a 95% credible interval. This is shown for mouse datasets 1, 2, and 3, and the ELSA human data. Variables are sorted by the ascending mean-timescale over all datasets. We see that there is a broad range of time-scales.

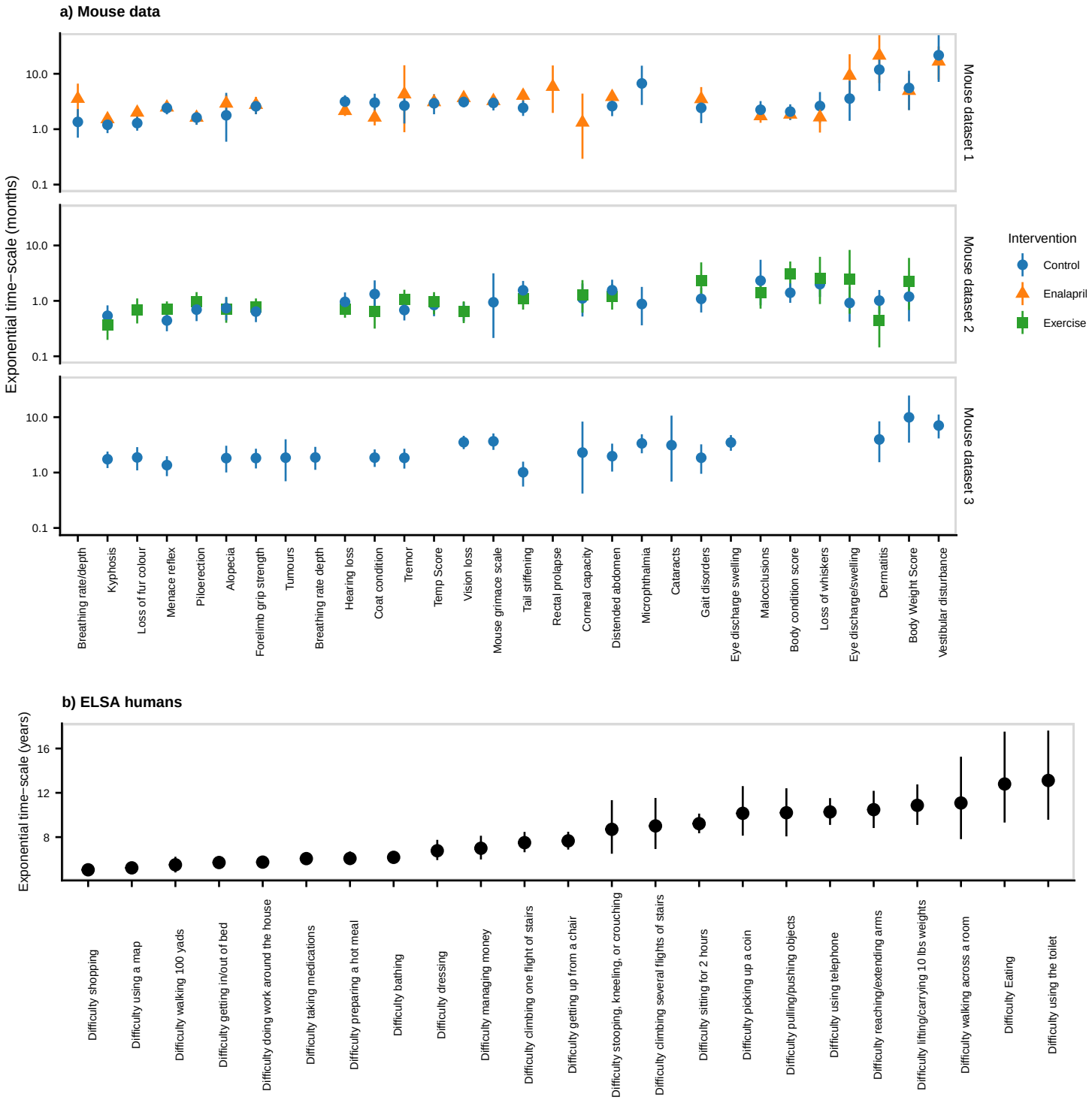


Figure 5-figure supplement 2. **Time-scales of robustness for individual deficits.** Exponential time-scales for undamaged state survival curves. Exponential time-scales for undamaged state survival curves. Exponential survival models are fit for the survival of the undamaged state for each binary health variable. Time-scales are computed as the inverse of the exponential rate, and the posterior median is shown as a point, with a 95% credible interval. This is shown for mouse datasets 1, 2, and 3, and the ELSA human data. Variables are sorted by the ascending mean-timescale over all datasets. We see that there is a broad range of time-scales.