The potential for complex computational models of aging

Spencer Farrell [a], Garrett Stubbings [a], Kenneth Rockwood [b], Arnold Mitnitski [b], Andrew Rutenberg [a],*

[a] *Department of Physics and Atmospheric Science, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4R2*
[b] *Division of Geriatric Medicine, Dalhousie University, Halifax, Nova Scotia, Canada B3H 2E1*

A B S T R A C T

The gradual accumulation of damage and dysregulation during the aging of living organisms can be quantified. Even so, the aging process is complex and has multiple interacting physiological scales – from the molecular to cellular to whole tissues. In the face of this complexity, we can significantly advance our understanding of aging with the use of computational models that simulate realistic individual trajectories of health as well as mortality. To do so, they must be systems-level models that incorporate interactions between measurable aspects of age-associated changes. To incorporate individual variability in the aging process, models must be stochastic. To be useful they should also be predictive, and so must be fit or parameterized by data from large populations of aging individuals. In this perspective, we outline where we have been, where we are, and where we hope to go with such computational models of aging. Our focus is on data-driven systems-level models, and on their great potential in aging research.

1. Introduction: Challenges of studying aging

Computational models are essential to make state-of-the-art predictions or to understand mechanisms within complex non-linear, stochastic, and interconnected systems such as the economy, the weather, or the climate. In this section we outline how aging organisms also represent complex, interconnected dynamical systems that are challenging to study.

Aging populations exhibit increasing mortality rates. For humans, the risk of dying increases approximately exponentially for older ages – the famous Gompertz law of mortality (Kirkwood, 2015). Before death, individual health can be assessed and summarized in many ways. One such measure is provided by the Frailty Index (FI) which is the proportion of "things wrong" from a large selection of possible age-related deficits of health and function (Mitnitski et al., 2001). The FI is robust, flexible, and is strongly correlated with various outcome measures including mortality (Rockwood et al., 2005; Evans et al., 2014). Alternatively, Biological Age (BA) is an "effective age" defined in terms of an individual's health, often using molecular aspects of health such as epigenetic methylation (Hannum et al., 2013; Horvath, 2013; Levine, 2020). Other summary measures of health have been developed, including allostatic load (McEwen and Stellar, 1993), and physiological disregulation (Milot et al., 2014). Different summary measures of health

are not necessarily strongly correlated with each other at the individual level (Li et al., 2020), indicating that aging is a multi-dimensional process.

As assessed by the FI, the distribution of health measures broadens with age, corresponding to distinctive individual trajectories of health (Rockwood et al., 2004). Worsening health over an individual's life is a random process, and is described as a stochastic accumulation of damage. It is thought that this damage underlies the increased mortality with age that is characterized by Gompertz's law (Gavrilov and Gavrilova, 2001).

Remarkably, even relatively simple empirical observations about aging and mortality are not well understood. Resolving the question of whether Gompertz's law applies for extremely old populations or whether they exhibit a mortality deceleration or plateau remains challenging due to small data-sets (Gavrilov and Gavrilova, 2019). The 'mortality-morbidity paradox' whereby female populations live longer than male populations, despite male populations apparently having better health, remains largely unexplained (Gordon and Hubbard, 2018; Gordon et al., 2017; Kulminski et al., 2008). The mechanisms behind historically changing health and mortality within national populations (Crimmins, 2015; Colchero et al., 2016), or behind differences between different socio-economic groups within a population (Andrew et al., 2012), are difficult to assess.

---

* Corresponding author.
  E-mail addresses: spencer.farrell@dal.ca (S. Farrell), adr@dal.ca (A. Rutenberg).

We see four big questions in aging research: how can we better observe health and mortality across large populations, how can we better understand the mechanisms or causes underlying what we observe, how can we better predict outcomes at an individual or population level, and, finally, how can we better intervene to decrease mortality and to improve health during aging? The challenges implicit in addressing these questions are interconnected: progress in any of these directions will support and direct progress in the others.

Success in aging research crucially depends on the broad availability of high-quality data. National studies, especially those that include longitudinal data on study participants, such as the CSHA (Canadian Study of Health and Aging Working Group, 1994), CLSA (Raina et al., 2009), NHANES (Centers for Disease Control and Prevention National Center for Health Statistics, 2014), BLSA (Ferrucci, 2008), ELSA (Steptoe et al., 2014), and the UK Biobank (Sudlow et al., 2015), are of increasing importance and utility. Emerging sources of "big"-data include electronic health records (EHR) (Clegg et al., 2016), molecular 'omics, and individual telemetry provided by health monitors or cell-phones. How and what can we learn from these new sources of aging data?

There are also many scales of health measures to consider: from molecular and cellular, to tissue, to organismal – including functional or social aspects of the organism. For example, at the molecular scale methylation clocks have emerged as convenient epigenetic hallmarks of health and aging (Hannum et al., 2013; Horvath, 2013). Other high-throughput technologies such as genomics, transcriptomics, proteomics, metabolomics, and microbiomics provide ways of measuring large amounts of data in aging studies (Livshits et al., 2018; Lehallier et al., 2019; Ahadi et al., 2020). Conversely, clinically relevant aspects of health such as activities of daily living (ADL) and other measures of functional disability are particularly important to the aging of individual older adults. Such 'higher' levels of function are dependent on many aspects of 'lower-level' biological and molecular function in a variety of tissues. How can we relate and usefully combine observations at different scales?

Continuing the historical advances in either life expectancy or healthy-aging will be increasingly challenging but, naturally, is of great interest to the geroscience community. Targeted interventions for aging individuals will often be in the context of significant comorbidities or polypharmacy. Systemic treatments such as exercise (Fried, 2016; Partridge et al., 2018), caloric restriction (Most et al., 2017; Mattson et al., 2017), or senolytics (Xu et al., 2018), act at cellular or molecular scales but the desired effects are often at the organismal scale. Understanding how different scales of organismal function interact with each other as a dynamical system should help us to effectively translate advances from one scale to another, and identify interventions that target the lower-level biological function before it manifests as functional disability (Ferrucci et al., 2018). Optimal interventions will likely depend on the health-state and age of the individual. How can we best provide individualized medicine for aging individuals?

Animal models of aging have instructive similarities and differences with respect to human aging (Cohen, 2018). Furthermore, efficient automated image analysis is starting to lead to high-quality longitudinal studies of model organisms (see e.g., worms Swierczek et al., 2011; Zhang et al., 2016, or flies Seroude et al., 2002). Developing insights from frequently-measured high-dimensional organismal health states within large model populations will require new suites of analysis and modelling tools, but should yield deeper understanding of the aging process. Simple animal models are particularly amenable to studying the multi-scale effects of controlled interventions.

Given the complexity of the aging process, how can computational models make use of available and emerging sources of data in order to improve our understanding of aging within and between populations or species, to better predict individual aging outcomes, and to both understand existing interventions and to develop better individualized interventions in the aging process? Table 1 summarizes some of the

Table 1
Some challenges and promise of complex computational models of aging

Challenges	Promise
Using large heterogeneous and sparse longitudinal datasets	Better individualized predictions
	(using E-health and self-reported data)
Making population context explicit	Better population health
	Better synthetic populations
Including multiple scales	Better translation of lab-based interventions
(e.g., cellular to functional) Understanding mechanisms	
Predicting and understanding effects of health interventions	Personalized treatment
	Improved treatment of comorbidities

particular challenges facing us in developing computational models of aging organisms, together with the potential benefits of meeting these challenges. We will return to these challenges in more detail at the end of this perspective.

2. Theoretical approaches to aging

Conceptual models such as the hallmarks of aging (López-Otín et al., 2013), seven pillars of aging (Kennedy et al., 2014), or damage accumulation (Kirkwood, 2005) can provide powerful frameworks for discussion or interpretation of quantitative results, but they are not typically quantitative themselves. While they organize how we think about aging, they do not directly help us to quantitatively characterize observed data or to make quantitative predictions.

The lightest quantitative models are largely descriptive, or phenomenological. An example is Gompertz's exponential increase of mortality with age for older adults (Kirkwood, 2015). This is not an exact model over any age range, but it is a useful approximation. Related to that is the Strehler-Mildvan correlation between the amplitude and exponent of the Gompertz law (Tarkhov et al., 2017), or the discussion of maximum human lifespan (Dolgin, 2018). Such phenomenological models beg explanation of mechanisms, but typically do not provide or require that explanation themselves. Furthermore, such simple quantitative models are not able to describe multiple organismal scales of aging. For example, they describe mortality but not health. As a result, they are not typically useful for predicting individual health outcomes.

Predictive models of aging make stronger quantitative assumptions. An example is the proportional hazards model (Cox, 1972). A review of mortality models is provided by Yashin et al. (2000). Yashin and colleagues also developed the quadratic hazards model (or stochastic process model, SPM) of aging (Yashin et al., 2012; Arbeev et al., 2016). It assumes that individual deviations from age-dependent norms of physiological measures interact with and exacerbate each other. The SPM is an example of a dynamical model of aging.

2.1. Dynamical models of aging

Dynamical models can be used to explicitly simulate individual health trajectories vs. age, i.e. longitudinal data. Because they can generate synthetic individual health data, they can serve much the same role as model organisms – whereby differences with respect to human aging can be significant but are hopefully also informative.

Dynamical models can build in explicit interpretable "mechanisms". These do not necessarily need to be fundamental biological mechanisms, but can be "effective" mechanisms relating variables that interact indirectly. Any success of a dynamical model in reproducing real-world phenomena then suggests the viability of these underlying mechanisms, and allows the modeller to explore other phenomena that arise from the same mechanisms. While such mechanisms cannot be immediately taken as real, the model can be used to identify further

experiments and data to test predictions arising from the mechanisms.

The concept of interactions between health measures can be embodied in a network of pair-wise interactions, where nodes (or "vertices") are the health measures while the interactions are connections between nodes (links, or "edges"). These interactions can be general, and not just the multiplicative interactions usually considered in regression models. Such explicit networks were already used to model mortality (Gavrilov and Gavrilova, 2001; Vural et al., 2014). We developed what we now call a generic network model (GNM) to also model health measures such as the FI, and found that we could describe both population-level aging and mortality with a simple network of interactions that could be implemented on a computer to generate large synthetic populations (Taneja et al., 2016; Farrell et al., 2016, 2018).

The health attributes (nodes) of our GNM did not directly correspond to specific observed health attributes. The reason for that approach was simplicity: all connected nodes have similar interactions, with simple undirected connections of equal weight, as illustrated in Fig. 1. This enabled us to capture population-level health and mortality with only a few parameters. Such generic networked models are useful for conceptual explorations of aging, and have been used to understand mortality curves in different species (Vural et al., 2014; Stroustrup et al., 2016), to explore the mutual information between health states and mortality (Farrell et al., 2018), and to explore how to increase longevity through optimized maintenance (Sun et al., 2020). Nevertheless, they are not able to capture realistic individual health states or model their detailed trajectories with age.

To be able to predict detailed individual health states, we would need to empirically capture the many distinct interactions between observed individual health attributes within a population. Reconstructing interactions from observed data is a daunting prospect. If hundreds of individual health attributes are measured, then there are tens of thousands of interactions to determine between all possible pairs of attributes. Initial progress is already being made on smaller-sized problems. Using only cross-sectional binarized health data, we have developed a network model that includes specific observed health attributes (Farrell et al., 2020). To accomplish this, the model parameters are distinct for each node and the network connections between nodes are distinctly weighted. An illustration of such a "weighted" network model (WNM) is shown in Fig. 2 for 10 nodes.

The WNM can be used to generate synthetic individual health trajectories until mortality from any starting point. For example, in Fig. 3 we illustrate the joint distribution of the FI at death and the death age for four individuals with specific baseline deficits at their baseline ages (either 65 or 85 years). We see how both baseline age and deficits affect both the death age and the overall health at death. Intriguingly, the most likely FI at death for the older individuals is 1, which is well above typically reported $FI_{max} \approx 0.7$ (Stubbings et al., 2020). Since we use only 10 deficits, the observed FI_{max} of this data is 1, but this does not explain the peak of the distribution at 1. The FI at death for individuals in this data is not observed, so this may be a prediction of a real effect related to terminal decline (Stolz et al., 2020) – however it may also partially be an artefact of over-simplified damage-rates within the model for individuals at higher FI (see discussion in Farrell et al. (2020)).

We can also look more closely at how health evolved up until death. These frailty trajectories are shown in Fig. 4, for individuals who died at their median predicted death age. This illustrates how baseline individual health and age can strongly affect subsequent health trajectories, and also how variable those trajectories are.
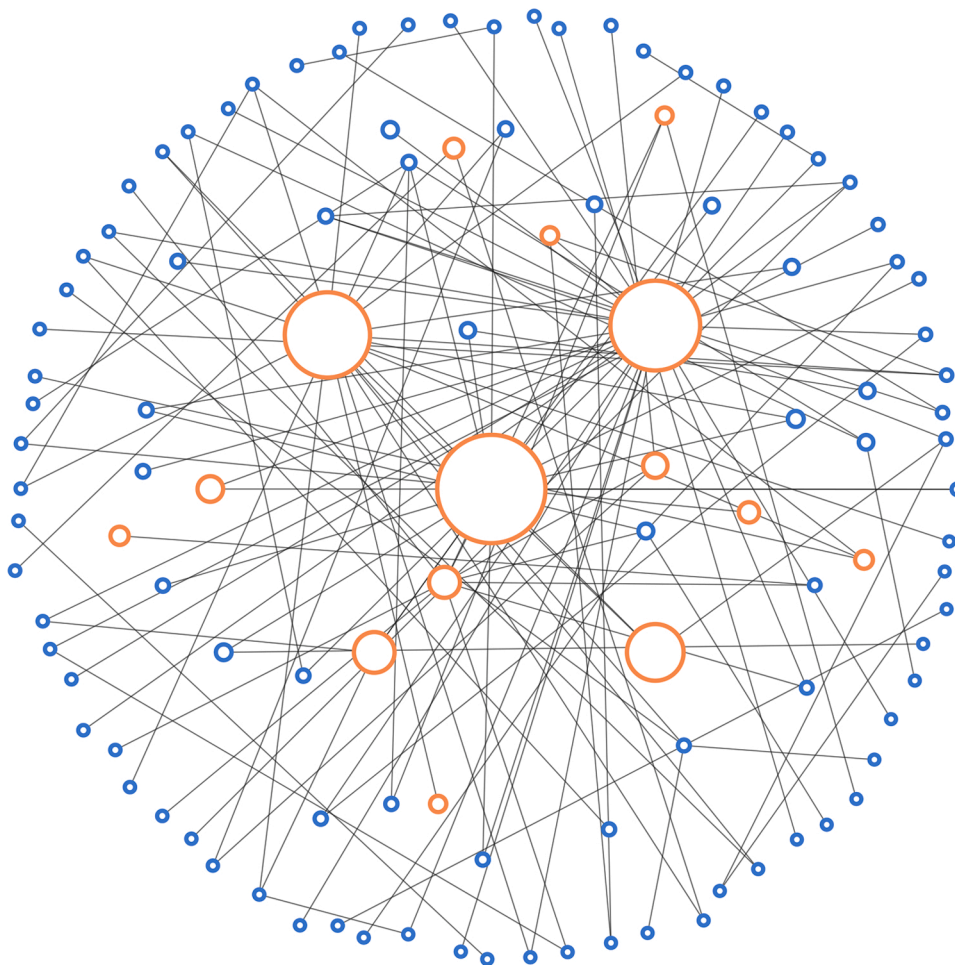
Fig. 1. Links (grey lines) between nodes (circles) in a generic network model (GNM) of aging. Shown are a selection of 112 nodes, out of 10^4 used in the model. Nodes do not represent particular observed health variables, but larger circles indicate nodes with more links. The most connected nodes are coloured orange, and are used in the aggregate health measure (frailty index, or FI) of the GNM. Note that the links do not change with age, while the binary health-states of the nodes do change with age between healthy and unhealthy. This model, with stochastic damage rates that can be implemented computationally (Taneja et al., 2016; Farrell et al., 2016, 2018), can generate large synthetic populations that reflect observed population-average age-dependent health and mortality rates.(For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
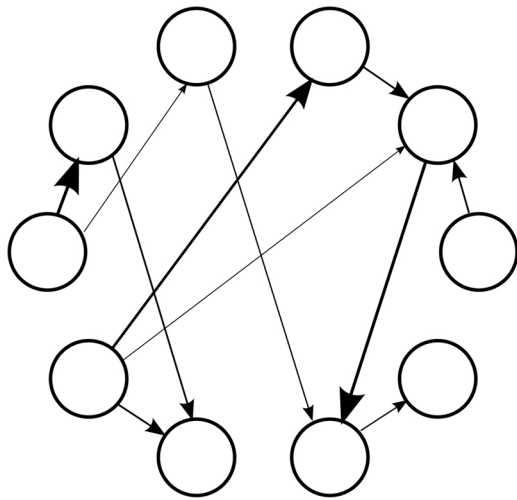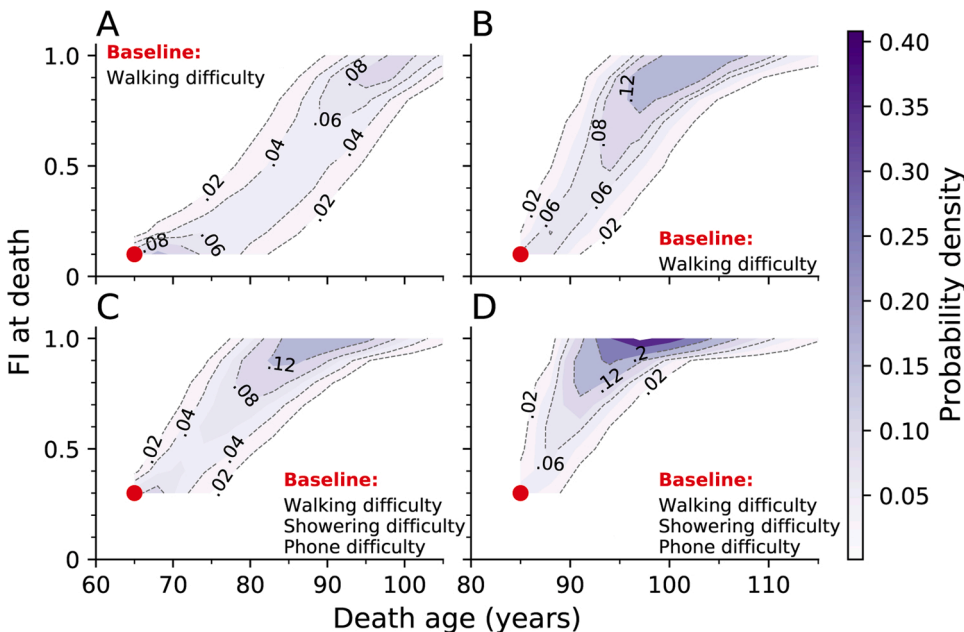
Fig. 2. Weighted and directed links (arrows) between nodes (circles) in a weighted network model (WNM) of aging. Shown are 10 nodes, each of which represents specific observed binary health variables from cross-sectional studies. Only the more significant links are shown, with weights represented by the line thickness and arrow size. This model, with stochastic damage rates that can be implemented computationally (Farrell et al., 2020), can generate large synthetic populations that reflect observed high-dimensional health-states and mortality over the populations used to fit the model. See Figs. 3 and 4 below. Note that the full model only has 10 nodes, representing just the observed health-states.

While these synthetic populations resemble observed aging human populations, we are not yet able to robustly infer specific interactions between the observed health attributes. Instead, we find that many different networks are consistent with the observed data. We are currently developing a more generalized approach, using continuous-valued longitudinal datasets, to make individual predictions of aging trajectories and to infer a robust network of interactions on the level of blood biomarkers and functional disabilities.

3. Computational models

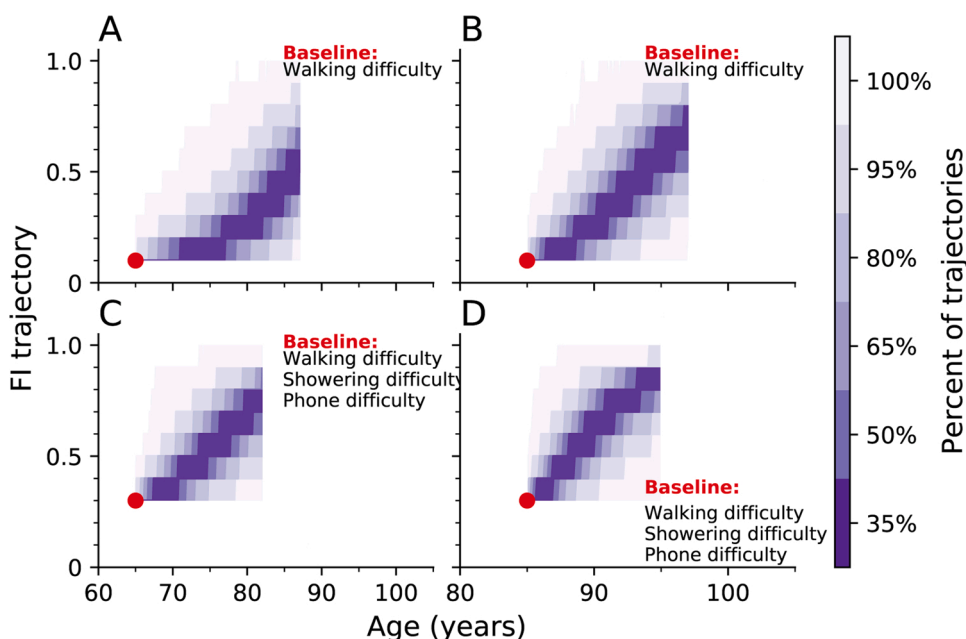All but the simplest models need to be implemented computationally. Computational models allow us to simulate and explore the quantitative consequences of various hypotheses. That is, since computational models require well-defined algorithms, they force us to make our assumptions explicitly. By varying those assumptions we can explore their consequences. This can clarify and illuminate possible mechanisms of aging.

Computational dynamical models can generate large synthetic populations of individuals with complete health trajectories and mortality. Such "perfect" data facilitates the systematic development of data analysis tools, including determining their statistical power for finite populations with missing data. More fundamentally, computational models also allow for a close examination of mechanisms at the population level: how does changing an assumption or a model parameter change the resulting health and mortality statistics of the population? The mechanisms behind the observed statistics of aging populations is a fundamental question of aging that can be answered, at least within the context of these model populations.

The other side of the coin is that a significant *disadvantage* of modelling is that every assumption needs to be explicitly built into the mathematical framework of the model, and every parameter needs to be determined. This is in contrast to observational studies of human or organismal populations, in which the assumptions and effective parameters are all implicitly included by biology. As a result, while natural observation naturally includes everything (including the "kitchen sink"), a modelling approach typically only builds a minimal framework to address the mechanisms under consideration. Identifying an appropriate minimal framework, as well as specific mechanisms to consider, can benefit from qualitative approaches such as group model building (Uleman et al., 2020). Such approaches can help focus modelling efforts towards specific achievable goals, and will be particularly useful where effects across multiple scales are modelled (see Kenzie et al., 2017) or more generally when faced with limited data.

The constructive flavour of modelling lends itself to inter-organismal comparisons, since we can ask whether parameter tuning alone can explain differences between organisms or whether the structure of the aging model needs also to be changed. Similar comparisons can be made between any distinct subpopulations of any one organism, including medically treated vs. untreated populations or genetically distinct populations.

Structural changes in models should not be needed to accommodate small differences between populations. Conversely, structural

Fig. 3. Simulated joint distributions of FI at death and death age from the network model described in Farrell et al. (2020), and illustrated in Fig. 2, given four individuals with the indicated initial baseline deficits (filled red points indicate their baseline age and FI). This demonstrates the capability of simulating populations of synthetic individuals, starting from different baseline conditions. The two columns show baseline ages of 65 (A and C) and 85 (B and D), the rows show different baseline FIs of 0.1 (A and B) and 0.3 (C and D), for an FI with 10 deficits.

Fig. 4. Intervals of sampled stochastic FI trajectories vs. age from the network model described in Farrell et al. (2020) and illustrated in Fig. 2, given given four individuals with the indicated initial baseline deficits. Trajectories are shown for samples with the median predicted death age, for the same individuals shown in Fig. 3. The purple shading indicates the percent of trajectories contained within the boundary of each region, similar to a confidence interval. The red circles indicates the baseline age and FI from which the simulation is started. The two columns show baseline ages of 65 (A and C) and 85 (B and D), the rows show different baseline FIs of 0.1 (A and B) and 0.3 (C and D), for a FI with 10 deficits.

differences between models will lead to distinctive effects that can be observed and therefore tested in observations of natural populations. However, testability will be challenging for mechanisms that are not already well characterized. For example, we expect that the effects of genetically heterogeneous human populations will eventually be important to characterize and include in models, but it will be hard to separate those effects from the intrinsic variability of the aging process. Nevertheless, a successful modelling framework should allow us to identify the statistical signatures of proposed mechanisms – which will facilitate subsequent testing.

To paraphrase Box (1976), all models are at least partially wrong but some can nevertheless be useful. However, rather than just trying to be mostly right or fairly useful, models should be improvable. This requires cycles of testing, development, and application to continually confront models with observable data. The benefit of this approach is that we can continually adjust our implicit or explicit assumptions to better and more usefully reflect emerging datasets.

3.1. Generalizability

Given the complexity of any organism, together with the complexity of the aging process, we can anticipate an enormous number of parameters required to tune complex models to fit a population. This tuning (also variously called fitting, learning, calibration, parameterization, or regression) is necessary if we want to do more than explore the qualitative consequences of a small set of model assumptions. Fitting the model to the data is necessary to make predictions for individuals, to compare populations, and to generate realistic synthetic populations.

We can distinguish between effective and fundamental parameters. Fundamental parameters are model independent, can be measured or derived with a variety of techniques, and are unchanged in different contexts. Effective parameters (like effective theories (Transtrum et al., 2015)), on the other hand, cannot be precisely replicated in different contexts and cannot be derived from any fundamental assumptions. We expect that almost all parameters of models of aging will be effective parameters, i.e. they will be at least somewhat dependent on the choice of model. Nevertheless, effective parameters should not be treated as arbitrary tuning-knobs of a model. More useful parameters will change less between studies, will be more interpretable, and will lead to better model predictions. Determining good model structure that facilitates useful parameterizations is an iterative process that is one goal of

successful modelling.

The technical details of fitting a model are well understood. Simple models can be hand-tuned to agree with population-level measures; however such simple models will not provide the best individual-level predictions. More sophisticated models can be fit with maximum likelihood, or other objective or "loss" functions, to obtain a model that best fits the available data. Bayesian approaches are also possible, where posterior distributions of parameters are obtained, rather than single point estimates.

An ideal data set would have a large homogeneous population, with complete, detailed health information that is longitudinally sampled frequently over individual lifetimes with uncensored mortality data. A computational model should be able to capture the important behavior exhibited in such data, so that it can then be used for individual predictions. Better computational models would provide better predictions. Are there other ways of distinguishing between such models? After all, even the best human data sets have small heterogeneous populations compared to national or global scales, with significant amounts of missing and censored data, and with irregular and infrequent longitudinal sampling with respect to the daily or weekly variability of our individual health status.

The answer hinges on how generalizable the model is to different data sets. If the model fit to one dataset poorly generalizes to other datasets – then the model has "failed". It has failed in a useful way (Box, 1976) if we can expand the model in an interpretable way to accommodate both datasets. It has failed in a disappointing way if we cannot, and if we cannot understand why not.

Ideally, any successful model would be tested with as many different datasets as possible. This builds confidence in the accuracy of the model, but conversely can uncover weaknesses in the model – suggesting areas to improve. A significant limiting factor currently is the lack of multiple large, publically available, longitudinal datasets with similar health observations.

It will be exciting to ask whether we can also generalize models to different applications. For example, consider interventions in the health of individual organisms due to drugs, surgery, treatment, lifestyle, accidents, illness, or (in the case of model organisms) experimental manipulation. Can a model predict the outcomes of such interventions? Better predictions could be used to both improve individual treatment plans and manage the health-care of increasingly greying populations (Harper, 2014).

Nevertheless, we might expect that any predictive model that is optimally tuned to predicting "natural" mortality or health outcomes will not be easily generalized to predict the outcomes of specific interventions. While part of this limitation naturally arises from the data used to train the model, some of this limitation will also come from the model structure itself, since computational models will not be able to accurately capture effects that are not allowed for in the model structure. Since the model structure itself may limit generalizability, some model structures ('types of models') will be better than others for this purpose.

3.2. Underdetermined parameters, overfitting, and bias

Any computational model of aging will also be at risk of underdetermined parameters, overfitting, and bias from the data sets used in training. These are generic problems of complex models.

Underdetermined parameters are parameters of the model that are not well constrained by the available data, but are nevertheless important for model functioning. A loose analogy is that while *a* hand may be needed for handwriting, *which* hand is much less constrained. Given limited data, strong correlations often exist between parameters that each can range widely in magnitude – this is the concept of *sloppiness* (Gutenkunst et al., 2007; Transtrum et al., 2015). As a result, when considered individually these poorly constrained or "sloppy" parameters have large uncertainties – even when the model can still make robust predictions.

To resolve this problem, we can focus on the predictions of complex models taking into account the uncertainties in underdetermined parameters, rather than the values of specific underdetermined parameters themselves (Gutenkunst et al., 2007). However, if direct interpretation of parameters is desired, determining specific parameter values can involve acquiring additional measurements, adding assumptions to the model, or otherwise improving model *identifiability* (Chis et al., 2016). Large clean datasets generated by computational models can be used to determine the types of observational data that would be needed to determine desired parameters.

Overfitting is another generic problem with complex models with many parameters. Here, parameter values are fine-tuned to extract small improvements in fitting to the training data at the expense of good performance with new data. Overfitting is assessed by using dedicated training data and separate but comparable test data to assess model performance. Since fitting typically occurs through an iterative computational algorithm, overfitting can be minimized by simply stopping the fitting process when model performance on a held-out portion of the training data (the validation or development data) begins to decline.

When test data is not comparable to training data then poor model performance can reflect poor generalizability due to limitations of the training data rather than due to overfitting. Some problem of generalizability arises in most training datasets, because they have biases in demographics (age distribution, sex, race), health-state of enrolled participants, medical treatment during the course of the study, or in any other possible category within the dataset.

Modelling bias can also arise due to the structure of model not being able to account for all aspects of the data. For example, survivor bias (Murphy et al., 2011) can be troublesome for models that do not capture mortality properly. When measured covariates have associations with mortality, the drop-out of individuals during the study due to mortality can bias the results. Models must account for survival effects, for example with joint longitudinal-survival models (Hickey et al., 2016) that model health and survival together – otherwise modelling efforts can erroneously try to accommodate survival effects within the disease progression itself.

3.3. Specific computational approaches

There are many possible approaches towards computational modelling. The most productive approaches will be determined by a combination of the background of the researcher, the problem at hand, and the data available. Agent-based modelling is one popular approach to managing multiple spatial and temporal scales in ecological systems (Grimm, 2005) or in socially interacting populations (Bonabeau, 2002). However, given the large heterogeneous data-sets and complexity inherent in aging organisms, we feel that Machine learning (ML) techniques are particularly promising for aging research.

4. Machine learning

Machine learning (ML) is a loosely-defined term for a collection of data-based models that are typically fit or "trained" with large high-dimensional data sets. Typical goals of ML approaches are classification (the most common application (Domingos, 2012), though not our focus here), regression, and generating synthetic samples with the same properties as the observed data.

Neural networks are often used in more sophisticated ML models (Goodfellow et al., 2016), as in deep learning (LeCun et al., 2015). Neural networks consist of layers of artificial neurons that each have many linearly combined input connections from previous layers, and many output connections to subsequent layers. All connection parameters for every neuron are trainable. Non-linear transformations in each neuron allow multiple layers (i.e. "deep" networks) to represent functions or relationships of arbitrary complexity (Leshno et al., 1993; Raghu et al., 2017).

Powerful neural networks have enormous numbers of parameters that must be trained for the network to represent a desired function. Neural networks are designed so that this training is computationally efficient. Overfitting can be a concern with so many parameters, and it is managed by careful use of regularization, which imposes restrictions on the parameters learned by the model. Test data, not used in training, is an important part of evaluating model performance and behavior.

Any unknown component of a model of aging can therefore be learned with a neural network, given sufficient training data. However, the researcher still needs to develop the overall structure of the model (i. e. how all the pieces glue together), choose appropriate neural network architectures, and manage the algorithms (and their "hyper"-parameters) that train the model while avoiding overfitting.

ML is rapidly developing, and new ML tools are easily learned and used after some expertise is gained. It is more challenging to achieve the goals of generalizability, where models perform well on data that is unlike the training set, and interpretability, where the mechanisms of the model can be understood and related to mechanisms in other model systems. Both of these goals are difficult because the flexibility of complex models required to achieve generalizability limits interpretability. Large data sets can help with generalizability by allowing the use of more complex models, but interpretability is an ongoing challenge (Rudin, 2019).

One strategy towards improving interpretability with ML is to develop models where knowledge of specific aging mechanisms are built into the model, while unknown components are learned with deep neural networks (using, e.g., differential equations to capture their behavior (Rackauckas et al., 2020)). With this approach, encoding the knowledge we already have about the aging system effectively constrains the more general ML approach. This maintains the flexibility that deep neural networks have to offer, allows the model to be trained with less data than a more general model, and adds interpretability to the model.

Nevertheless, caution will be needed in accepting novel mechanisms just because they lead to better predictions or model performance. While we do expect that correct mechanisms will lead to better model behavior and generalizability, we also expect that powerful ML approaches may

be able to perform reasonably well in spite of incorrect mechanisms. A critical approach will be called for, using data and predictions that can discriminate between putative mechanisms.

4.1. ML approaches in ageing research

Several ML models have been developed specifically for aging. Pierson et al. (2019) developed a model that infers rates of aging for individuals that correlate with risk factors of aging, and that can be used to forecast future health. Avchaciov et al. (2020) developed a model that describes the aging of mice with an inferred dynamical frailty index, which correlates with both mortality and treatment effects.

Similarly, machine learning has already had success in the estimation of biological age (Hannum et al., 2013; Horvath, 2013; Levine et al., 2018; Lu et al., 2019). With these models, many biological variables (e. g., DNA methylation levels) are reduced to a single estimate of biological age, which is found to be predictive of other health outcomes and mortality. The models used here are generally regularized linear models, due to the huge number of variables compared to the limited amount of data. As more data becomes available, more sophisticated techniques can be used for assessing biological age (Putin et al., 2016; Pyrkov et al., 2018; Schultz et al., 2019; Zhavoronkov and Mamoshina, 2019).

Nevertheless, models of biological age are not dynamical models – they cannot simulate the future health trajectories of individuals, but only summarize and interpret the current health state. Furthermore, since reducing health to a single variable cannot capture multi-dimensional aspects of health, we believe that developing dynamical models that address longitudinal trajectories across multiple health dimensions is a promising direction for machine learning in aging research.

To forecast multi-dimensional health trajectories, existing machine learning approaches for modelling disease progression could be adapted to model aging progression. While many of these do not model mortality (Schulam and Suchi, 2015; Alaa and van der Schaar, 2018; Fisher et al., 2019; Walsh et al., 2020), joint longitudinal-survival models could be adapted for this purpose (Lim and van der Schaar, 2018). A stochastic process model of aging has already been developed that models both health trajectories and mortality (Yashin et al., 2007, 2012; Arbeev et al., 2011, 2014), but it has not yet been applied to high-dimensional datasets.

Given efficient algorithms for parameter determination (learning) together with flexible functional dependence (deep learning), we see great promise for ML approaches in the study of aging. Natural applications are filling in missing data, identifying natural subpopulations or categories of aging organisms, incorporating multiple heterogeneous data sources, and modelling the aging process itself as a stochastic dynamical process.

In Table 2 we have listed some current work in the machine learning literature that we believe could be useful in applying machine learning to aging. While additional development of any existing techniques would be required for any specific problem, these approaches (and the references they cite) capture many useful ideas.

5. The challenges and promise of aging models

In the introduction, we listed four challenges of studying aging: how to better observe health in aging populations, how to better understand the mechanisms behind what we observe, how to better predict individual health, and how to better intervene in the aging process. We then highlighted particular challenges and promise of generalized computational models of aging in Table 1. Here, we provide more detail about the near-term challenges facing aging models – together with some of the opportunities that make facing these challenges worthwhile.

5.1. E-health, self-reported, and longitudinal data

Focused population surveys are expensive. Large scale studies such as ELSA (Steptoe et al., 2014) or NHANES (Centers for Disease Control and Prevention National Center for Health Statistics, 2014) are limited to populations on the order of 10000 individuals. Even the impressive UK Biobank has less than one million individuals (Sudlow et al., 2015). In contrast, the use of electronic health records (EHR) (Clegg et al., 2016) could eventually reach large fractions of national populations with lifetime longitudinal data. EHRs are therefore an attractive source of data on the aging process. Similarly, individual health tracking through e.g., smart watches, or through self-reporting, could also reach large fractions of national populations. These developments will provide natural datasets with large populations that have lifetime longitudinal information.

Significant biases are found in EHR data (Vassy et al., 2018) and also in self-reported health data (Zajacova and Dowd, 2011; Gunasekara et al., 2012). It will be difficult to explicitly account for these biases in order to reconcile EHR and self-reported data with corresponding national prospective studies from similar populations. Nevertheless, an immediate opportunity is to use this data for personalized health models and predictions.

More generally, large-scale longitudinal data collection provides an opportunity for aging models to employ these data to better capture the aging process of individuals, including individual variability. Computational models of aging are well placed to make use of longitudinal data given the vast amount of data potentially available. Natural questions include how much is gained by more frequent measurements, how to best handle variables observed at irregular time intervals and with varying degrees of missing observations, how to model mixtures of qualitative and quantitative measurements or of self-reported and molecular measures, and how to include individual health histories in individual health predictions.

5.2. Defining and comparing populations

Some über-model of aging might explicitly capture each aspect of individual variability, including a life-history of diet, lifestyle, injury, medication, and health-care. More realistically, most variability will first need to be captured implicitly within aging models through parameterization or model structure – tuned for different natural subpopulations. Race (Williams, 2005), sex (Gordon and Hubbard, 2018; Gordon et al., 2017), socioeconomic position (Knesebeck Ovd et al., 2007), social vulnerability (Wallace et al., 2015), access to health-care (Santana, 2000) or pensions (Aguila et al., 2018), rural/urban (Yu et al., 2012), and nationality, are all categories that have been studied by aging researchers. Chronic disease, genetic disorders, and certain patterns of multimorbidity or polypharmacy could also serve as natural categories. A challenge will be to reduce the significance of these natural categories for individuals by making the more of the implicit differences between the populations explicit – allowing for better individualized study of aging and treatment. To be able to achieve this requires good data coverage across many subpopulations, but also good models that

Table 2
Promising approaches for the modelling of aging.

Longitudinal data	
Transitions between disease states at discrete times	Liu et al. (2015), Alaa and van der Schaar (2018) Fisher et al. (2019), Walsh et al. (2020)
Irregularly timed observations	Rubanova et al. (2019), Schulam and Suchi (2015)
Understanding covariate effects and interactions	Timonen et al. (2020)
Interpretable latent variables	Avchaciov et al. (2020), Pierson et al. (2019)
Joint Models (survival)	Lim and van der Schaar (2018)

can characterize and model the differences.

Models of individual health with explicit context could then be used to generate synthetic populations that match measured or projected demographic information. This would be particularly useful for detailed projections of the effects of aging in population health.

5.3. Multiple scales and subsystems

Different physiological scales present exciting opportunities in aging research. For example, molecular data is appealing because it can be high throughput and low-bias. Nevertheless, outcomes at higher (functional) scales are typically of greater individual interest. One challenge is to identify interactions between scales, from molecular to behavioral, and to incorporate them in aging models (Ferrucci et al., 2018; Mitnitski and Rockwood, 2019; Kuo et al., 2020). Reliably bridging the scales, particular in light of patchwork individual data (over scale, over time, and over individual measures), is an important challenge. Understanding how different scales work is the essence of understanding the aging process. How does damage propagate from the molecular to activities of daily living? Conversely, how do interventions of lifestyle or injuries propagate towards the molecular?

Diseases (such as Alzheimer's (Fisher et al., 2019)), tissues (such as the brain (Daunizeau et al., 2011)), or biological subfields (such as systems biology (Dada and Mendes, 2011)) each have distinct data-sets and modelling approaches. Particular aspects of the aging process, such as cellular senescence (Karin et al., 2019), can be similarly detailed. We see two promising ways to combine specialized approaches with more generalized models of organismal aging. The first is to identify key summary measures from detailed models, and to train generalized models with accordingly pre-processed data. The second is to include generalized models as background aging processes within the more specialized models. Both should be explored, so that aging processes are more routinely combined with emerging biological, physiological, and medical models.

5.4. Predictions, treatment, and interventions

Although individual predictions of health trajectories and mortality are natural goals for computational models of aging, a challenge is how to evaluate and judge the quality of the predictions, given the variety of different studies and possible outcomes. Evaluation of predictive quality is straight-forward retrospectively within the same dataset, by using separate training and test populations with either cross-sectional or longitudinal data with linked mortality. The determination of what quality of predictions are possible with what sort of data for individuals of a given age and health status will be important questions to answer.

Being able to predict the results of medical interventions (including medication), of illness or injury, or of life-style interventions such as exercise for aging individuals would be game-changing since it could be used to improve personalized treatment. Most individuals experience many such interventions over their lifetime, so these are implicitly and approximately included in models of national aging populations. Indeed, we assume that many such interventions are the origin of most national differences or differences within a national population over history.

A grand challenge will be to make many of these interventions explicit, particularly within models of individual health during aging. If successful, such explicit models will allow better individual prediction, better identification of intrinsic variability, and the ability to tailor or individualize interventions to better reflect individual priorities. To do this well we may need to include earlier data across individual life courses for large populations, including electronic health records and other longitudinal data.

Current research is often focused on the "diseases of aging", such as Alzheimer's disease, cancer, and cardiovascular disease. While we are optimistic that complex computational models of aging can be applied to

these conditions, doing so will require large-scale data for long-term outcomes of these specific diseases and possible interventions – which may be difficult to obtain.

In the short term, there is a use for dynamical models in clinical trials since models can be used to generate large synthetic control populations (Fisher et al., 2019; Walsh et al., 2020). Models could synthetically create control arms that are better matched in age, sex, and original health status with respect to the treatment arm. Computational models may also be useful to explore and understand the effects of comorbidities and their treatments.

6. Looking ahead

Early modelling has been restricted to simple theoretical or statistical explorations of the aging process, through damage accumulation or regression models. Though this approach has limited ability to predict individual health, it has advanced our conceptual understanding of how aging could work.

More recently, various models have started to address observational data that includes the detailed health and mortality of large numbers of individuals, which we call "networked" models since they capture interactions between different aspects of individual health. Our work in this area has included explicit complex networks, but the networks can also be theoretical (Yashin et al., 2012; Arbeev et al., 2016), correlational (Hidalgo et al., 2009; Roque et al., 2011; García-Peña et al., 2019), or implicit in the approach. Few models have addressed both individual health *and* mortality, though these are now starting to emerge (Farrell et al., 2020).

Once multiple models with both health and mortality are developed, then the natural scientific selection of "better" models can proceed by confronting their simulated results with observed data. Natural measures of goodness of models include predictive quality, generalizability across different population demographics (including age and health, but also sex and chronic conditions), interpretability, and the ability to effectively and efficiently train with big heterogeneous data sets. The ability to efficiently and effectively predict future individual health trajectories will be revolutionary, particularly if models include the effects of injury and disease, or the benefit of various medical and pharmacological interventions.

While observational data sets will only increase both in the number of individuals, in the number of physiological aspects of health that are reported, and in the frequency of longitudinal measurements, the amount of easy-available data available to train, test, and compare modelling approaches is still limited. Public "challenge" datasets could provide realistically imperfect but extensive longitudinal health data together with mortality statistics to allow for comparison between and improvement of modelling approaches. Providing raw data together with cleaned data is important, since improvements in data-cleaning (Van den Broeck et al., 2005) can also lead to model improvement – and computational pipelines of data-cleaning will be increasingly necessary for large population studies.

We will never achieve a "death-clock" where we can precisely predict an individual's death, nor a health-calendar of precisely how their health will change as they age. Nevertheless, we may be able to classify and identify useful aging phenotypes, to obtain good predictions of individual health-trajectories and mortality, and to identify the most useful health interventions for a given individual. Because computational models can capture the effects of many interacting aspects of human physiology, they are promising tools to use to help to address these questions.

How computational models can and will be used will depend on how successful they become. We believe that they will lead to a deeper understanding of how aging works, both for human aging and for model organisms. By incorporating many different mechanistic effects within and between different organismal scales, computational modelling could reach towards an overarching, contingent, and quantitative theory

of aging.

More prosaically, computational models could help to control for the effects of different populations, or to improve national or regional comparisons of the determinants of health. We also expect that models will be able to capture the effects of various health interventions at the individual level. If models become sufficiently good, they would be able to help individuals to develop and adapt their personal health plans. We are hopeful.

Acknowledgements

References

Aguila, E., López-Ortega, M., Robledo, L.M.G., 2018. Non-contributory pension programs and frailty of older adults: evidence from Mexico. PLoS ONE 13 (11), e0206792.

Ahadi, S., Zhou, W., Rose, S.M.S.F., Sailani, M.R., Contrepois, K., Avina, M., Ashland, M., Brunet, A., Snyder, M., 2020. Personal aging markers and ageotypes revealed by deep longitudinal profiling. Nat. Med. 26, 83–90.

Alaa AM, van der Schaar M (2018) Forecasting individualized disease trajectories using interpretable deep learning. arXiv arXiv:1810.10489v1.

Andrew, M.K., Mitnitski, A., Kirkland, S.A., Rockwood, K., 2012. The impact of social vulnerability on the survival of the fittest older adults. Age Ageing 41 (2), 161–165.

Arbeev, K.G., Ukraintseva, S., Akushevich, I., Kulminski, A.M., 2011. Age trajectories of physiological indices in relation to healthy life course. Mech. Ageing Dev. 132, 93–102.

Arbeev, K.G., Akushevich, I., Kulminski, A.M., Ukraintseva, S.V., Yashin, A.I., 2014. Joint analyses of longitudinal and time-to-event data in research on aging: implications for predicting health and survival. Front. Public Health 2 (228).

Arbeev, K.G., Ukraintseva, S.V., Yashin, A.I., 2016. Dynamics of biomarkers in relation to aging and mortality. Mech. Ageing Dev. 156, 42–54.

Avchaciov, K., Antoch, M.P., Andrianova, E.L., Tarkhov, A.E., Menshikov, L.I., Burmistrova, O., Gudkov, A.V., Fedichev, P.O., 2020. Identification of a blood test-based biomarker of aging through deep learning of aging trajectories in large phenotypic datasets of mice. bioRxiv. https://doi.org/10.1101/2020.01.23.917286.

Bonabeau, E., 2002. Agent-based modeling: Methods and techniques for simulating human systems. Proc. Natl. Acad. Sci. 99 (suppl 3), 7280–7287.

Box, G.E.P., 1976. Science and statistics. J Am Stat Assoc. 71, 791–799.

Uleman, J.F., Melis, R.J.F., Quax, R., van der Zee, E.A., Thijssen, D., Dresler, M., van de Rest, O., van der Velpen, I.F., Adams, H.H.H., Ben Schmand IMCMdK, de Bresser, J., Richard, E., Verbeek, M., Hoekstra, A.G., Rouwette, E.A.J.A., Rikkert, M.G.M.O., 2020. Mapping the multicausality of Alzheimer's disease through group model building. GeroScience.

Van den Broeck, J., Cunningham, S.A., Eeckels, R., Herbst, K., 2005. Data cleaning: detecting, diagnosing, and editing data abnormalities. PLoS Medicine 2 (10), e267.

Canadian Study of Health and Aging Working Group, 1994. Canadian study of health and aging: study methods and prevalence of dementia. Can. Med. Assoc. J. 150 (6), 899.

Centers for Disease Control and Prevention National Center for Health Statistics (Updated 2014) National health and nutrition examination survey data. http://www.cdc.gov/nchs/nhanes.htm.

Chis, O.T., Villaverde, A.F., Banga, J.R., Balsa-Canto, E., 2016. On the relationship between sloppiness and identifiability. Math. Biosci. 282, 147–161.

Clegg, A., Bates, C., Young, J., Ryan, R., Nichols, L., Teale, E.A., Mohammed, M.A., Parry, J., Marshall, T., 2016. Development and validation of an electronic frailty index using routine primary care electronic health record data. Age Ageing 45, 353–360.

Cohen, A.A., 2018. Aging across the tree of life: The importance of a comparative perspective for the use of animal models in aging. Biochim. Biophys. Acta 1864 (9), 2680–2689.

Colchero, F., Rau, R., Jones, O.R., Barthold, J.A., Conde, D.A., Lenart, A., Nemeth, L., Scheuerlein, A., Schoeley, J., Torres, C., Zarulli, V., Altmann, J., Brockman, D.K., Bronikowski, A.M., Fedigan, L.M., Pusey, A.E., Stoinski, T.S., Strier, K.B., Baudisch, A., Alberts, S.C., Vaupel, J.W., 2016. The emergence of longevous populations. Proc. Natl. Acad. Sci. 113 (48), E7681–E7690.

Cox, D.R., 1972. Regression models and life-tables. J. R Stat Soc Series B Stat Methodol 34 (2), 187–220.

Crimmins, E.M., 2015. Lifespan and healthspan: past, present, and promise. The Gerontologist 55 (6), 901–911.

Dada, J.O., Mendes, P., 2011. Multi-scale modelling and simulation in systems biology. Integr. Biol. 3 (2), 86.

Daunizeau, J., David, O., Stephan, K.E., 2011. Dynamic causal modelling - A critical review of the biophysical and statistical foundations. Neuroimage 58 (2), 312–322.

Dolgin, E., 2018. There's no limit to longevity, says study that revives human lifespan debate. Nature 559 (7712), 14–15.

Domingos, P., 2012. A few useful things to know about machine learning. CACM 55 (10), 78–87.

Evans, S.J., Sayers, M., Mitnitski, A., Rockwood, K., 2014. The risk of adverse outcomes in hospitalized older patients in relation to a frailty index based on a comprehensive geriatric assessment. Age Ageing 43 (1), 127–132.

Farrell, S.G., Mitnitski, A.B., Rockwood, K., Rutenberg, A.D., 2016. Network model of human aging: frailty limits and information measures. Phys. Rev. E. 94 (5), 052409.

Farrell, S.G., Mitnitski, A.B., Theou, O., Rockwood, K., Rutenberg, A.D., 2018. Probing the network structure of health deficits in human aging. Phys. Rev. E. 98 (3), 187.

Farrell, S.G., Mitnitski, A.B., Theou, O., Rockwood, K., Rutenberg, A.D., 2020. Generating individual aging trajectories with a network model using cross-sectional data. Scientific Reports 10, 19833. https://doi.org/10.1038/s41598-020-76827-3.

Ferrucci, L., 2008. The Baltimore Longitudinal Study of Aging (BLSA): a 50-year-long journey and plans for the future. J. Gerontol. A Biol. Sci. 63 (12), 1416–1419.

Ferrucci, L., Levine, M.E., Kuo, P.L., Simonsick, E.M., 2018. Time and the metrics of aging. Circ. Res. 123, 740–744.

Fisher, C.K., Smith, A.M., Walsh, J.R., 2019. Machine learning for comprehensive forecasting of Alzheimer' Disease progression. Sci. Rep. 9 (1), 1–14.

Fried, L.P., 2016. Interventions for human frailty: Physical activity as a model. Cold Spring Harb. Perspect. Med. 6 (6), a025916.

García-Peña, C., Ramírez-Aldana, R., Parra-Rodriguez, L., Gómez-Verján, J.C., Pérez-Zepeda, M.U., Gutiérrez-Robledo, L.M., 2019. Network analysis of frailty and aging: empirical data from the Mexican Health and Aging Study. Exp. Gerontol. 128, 110747.

Gavrilov, L.A., Gavrilova, N.S., 2001. The reliability theory of aging and longevity. J. Theor. Biol. 213 (4), 527–545.

Gavrilov, L.A., Gavrilova, N.S., 2019. New trend in old-age mortality: Gompertzialization of mortality trajectory. Gerontology 65 (5), 451–457.

Goodfellow I, Bengio Y, Courville A (2016) Deep Learning. MIT Press, http://www.deeplearningbook.org.

Gordon, E., Peel, N., Samanta, M., Theou, O., Howlett, S., Hubbard, R., 2017. Sex differences in frailty: a systematic review and meta-analysis. Exp. Gerontol. 89, 30–40.

Gordon, E.H., Hubbard, R.E., 2018. Physiological basis for sex differences in frailty. Curr. Opin Physio. 6, 10–15.

Grimm, V., 2005. Pattern-oriented modeling of agent-based complex systems: lessons from ecology. Science 310 (5750), 987–991.

Gunasekara, F.I., Carter, K., Blakely, T., 2012. Comparing self-rated health and self-assessed change in health in a longitudinal survey: which is more valid? Soc. Sci. Med. 74 (7), 1117–1124.

Gutenkunst, R.N., Waterfall, J.J., Casey, F.P., Brown, K.S., Myers, C.R., Sethna, J.P., 2007. Universally sloppy parameter sensitivities in systems biology models. PLoS Comput Biol 3, 10.

Hannum, G., Guinney, J., Zhao, L., Zhang, L., Hughes, G., Sadda, S., Klotzle, B., Bibikova, M., Fan, J.B., Gao, Y., Deconde, R., Chen, M., Rajapakse, I., Friend, S., Ideker, T., Zhang, K., 2013. Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol. Cell 49 (2), 359–367.

Harper, S., 2014. Economic and social implications of aging societies. Science 346 (6209), 587–591.

Hickey, G.L., Philipson, P., Jorgenson, A., Kolamunnage-Dona, 2016. Joint modelling of time-to-event and multivariate longitudinal outcomes: recent developments and issues. BMC Med. Res. Methodol. 16 (117).

Hidalgo, C.A., Blumm, N., Barabási, A.L., Christakis, N.A., 2009. A dynamic network approach for the study of human phenotypes. PLOS Comput. Biol. 5 (4), 1–11.

Horvath, S., 2013. DNA methylation age of human tissues and cell types. Genome Biol. 14, R115.

Karin, O., Agrawal, A., Porat, Z., Krizhanovsky, V., Alon, U., 2019. Senescent cell turnover slows with age providing an explanation for the Gompertz law. Nat. Commun. 10 (1), 1–9.

Kennedy, B.K., Berger, S.L., Brunet, A., Campisi, J., Cuervo, A.M., Epel, E.S., Franceschi, C., Lithgow, G.J., Morimoto, R.I., Pessin, J.E., Rando, T.A., Richardson, A., Schadt, E.E., Wyss-Coray, T., Sierra, F., 2014. Aging: a common driver of chronic diseases and a target for novel interventions. Cell 159 (4), 709–713.

Kenzie, E.S., Parks, E.L., Bigler, E.D., Lim, M.M., Chesnutt, J.C., Wakeland, W., 2017. Concussion as a multi-scale complex system: an interdisciplinary synthesis of current knowledge. Front. Neurol. 8, e160294.

Kirkwood, T.B.L., 2005. Understanding the odd science of aging. Cell 120, 437–447.

Kirkwood, T.B.L., 2015. Deciphering death: a commentary on Gompertz (1825) On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. Philos. Trans. R. Soc. 370 (1666), 20140379.

Knesebeck Ovd, Wahrendorf, M., Hyde, M., Siegrist, J., 2007. Socio-economic position and quality of life among older people in 10 European countries: results of the SHAREstudy. Ageing. Soc. 27 (2), 269–284.

Kulminski, A.M., Culminskaya, I.V., Ukraintseva, S.V., Arbeev, K.G., Land, K.C., Yashin, A.I., 2008. Sex-specific health deterioration and mortality: the morbidity–mortality paradox over age and time. Exp. Gerontol. 43 (12), 1052–1057.

Kuo, P.L., Schrack, J.A., Shardell, M.D., Levine, M., Moore, A.Z., An, Y., Elango, P., Karikkineth, A., Tanaka, T., de Cabo, R., Zukley, L.M., AlGhatrif, M., Chia, C.W., Simonsick, E.M., Egan, J.M., Resnick, S.M., Ferrucci, L., 2020. A roadmap to build a phenotypic metric of ageing: insights from the Baltimore Longitudinal Study of Aging. J. Intern. Med. 287 (4), 373–394.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436–444.

Lehallier, B., Gate, D., Schaum, N., Nanasi, T., Lee, S.E., Yousef, H., Losada, P.M., Berdnik, D., Keller, A., Verghese, J., Sathyan, S., Franceschi, C., Milman, S.,

Barzilai, N., Wyss-Coray, T., 2019. Undulating changes in human plasma proteome profiles across the lifespan. Nat. Med. 25, 1843–1850.

Leshno, M., Lin, V.Y., Pinkus, A., Schocken, S., 1993. Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. Neural. Netw. 6 (6), 861–867.

Levine, M.E., 2020. Assessment of epigenetic clocks as biomarkers of aging in basic and population research. J. Gerontol. A Biol. Sci. 75 (3), 463–465.

Levine, M.E., Lu, A.T., Quach, A., Chen, B.H., Assimes, T.L., Bandinelli, S., Hou, L., Baccarelli, A.A., Stewart, J.D., Li, Y., Whitsel, E.A., Wilson, J.G., Reiner, A.P., Aviv, A., Lohman, K., Liu, Y., Ferrucci, L., Horvath, S., 2018. An epigenetic biomarker of aging for lifespan and healthspan. Aging (Milano) 10 (4), 573–591.

Li, X., Ploner, A., Wang, Y., Magnusson, P.K., Reynolds, C., Finkel, D., Pedersen, N.L., Jylhävä, J., Hägg, S., 2020. Longitudinal trajectories, correlations and mortality associations of nine biological ages across 20-years follow-up. eLife 9, e51507.

Lim, B., van der Schaar, M., 2018. Disease-atlas: Navigating disease trajectories using deep learning. Proc. Mach. Learn. Res. 85, 137–160.

Liu, Y.Y., Li, S., Li, F., Song, L., Rehg, J.M., 2015. Efficient learning of continuous-time hidden Markov models for disease progression. Adv. Neural Inf. Process Syst. 3600–3608.

Livshits, G., Malkin, I., Bowyer, R.C.E., Verdi, S., Bell, J.T., Menni, C., Williams, F.M.K., Steves, C.J., 2018. Multi-omics analyses of frailty and chronic widespread musculoskeletal pain suggest involvement of shared neurological pathways. Pain 159 (12), 2565–2572.

López-Otín, C., Blasco, M.A., Partridge, L., Serrano, M., Kroemer, G., 2013. The hallmarks of aging. Cell 153, 1194–1217.

Lu, A.T., Quach, A., Wilson, J.G., Reiner, A.P., Aviv, A., Raj, K., Hou, L., Baccarelli, A.A., Li, Y., Stewart, J.D., Whitsel, E.A., Assimes, T.L., Ferrucci, L., Horvath, S., 2019. DNA methylation GrimAge strongly predicts lifespan and healthspan. Aging (Milano) 11 (2), 303–327.

Mattson, M.P., Longo, V.D., Harvie, M., 2017. Impact of intermittent fasting on health and disease processes. Ageing Res. Rev. 39, 46–58.

McEwen, B.S., Stellar, E., 1993. Stress and the individual: mechanisms leading to disease. Arch. Intern. Med. 153 (18), 2093–2101.

Milot, E., Morissette-Thomas, V., Li, Q., Fried, L.P., Ferrucci, L., Cohen, A.A., 2014. Trajectories of physiological dysregulation predicts mortality and health outcomes in a consistent manner across three populations. Mech. Ageing Dev. 141-142, 56–63.

Mitnitski, A., Rockwood, K., 2019. The Problem Of Integrating Biological And Clinical Markers Of Aging. In: Biomarkers of Human Aging. Healthy Ageing and Longevity, 10. Springer.

Mitnitski, A.B., Mogilner, A.J., Rockwood, K., 2001. Accumulation of deficits as a proxy measure of aging. Sci. World J. 1, 323–336.

Most, J., Tosti, V., MRedman, L., Fontana, L., 2017. Calorie restriction in humans: an update. Ageing Res. Rev. 39, 36–45.

Murphy, T.E., Han, L., Allore, H.G., Peduzzi, P.N., Gill, T.M., Lin, H., 2011. Treatment of death in the analysis of longitudinal studies of gerontological outcomes. J. Gerontol. A Biol. Sci. 66A (1), 109–114.

Partridge, L., Deelen, J., Slagboom, P.E., 2018. Facing up to the global challenges of ageing. Nat. Mach. Intell. 561 (7721), 45–56.

Pierson, E., Koh, P.W., Hashimoto, T., Koller, D., Liang, P., 2019. Inferring multidimensional rates of aging from cross-sectional data. Proc. Mach. Learn Res. 89, 97–107.

Putin, E., Mamoshina, P., Aliper, A., Korzinkin, M., Moskalev, A., Kolosov, A., Ostrovskiy, A., Cantor, C., Vijg, J., Zhavoronkov, A., 2016. Deep biomarkers of human aging: application of deep neural networks to biomarker development. Aging (Milano) 8 (5), 1021–1033.

Pyrkov, T.V., Slipensky, K., Barg, M., Kondrashin, A., Zhurov, B., Zenin, A., Pyatnitskiy, M., Menshikov, L., Markov, S., Fedichev, P.O., 2018. Extracting biological age from biomedical data via deep learning: too much of a good thing? Sci. Rep. 8 (5210).

Rackauckas, C., Ma, Y., Martensen, J., Warner, C., Zubov, K., Supekar, R., Skinner, D., Ramadhan, A., Edelman, A., 2020. Universal Differential Equations For Scientific Machine Learning arXiv 2001.04385.

Raghu, M., Poole, B., Kleinberg, J., Ganguli, S., Sohl-Dickstein, J., 2017. On the expressive power of deep neural networks. PMLR 70, 2847–2854.

Raina, P.S., Wolfson, C., Kirkland, S.A., Griffith, L.E., Oremus, M., Patterson, C., Tuokko, H., Penning, M., Balion, C.M., Hogan, D., Wister, A., Payette, H., Shannon, H., Brazil, K., 2009. The Canadian longitudinal study on aging (CLSA). Can. J. Aging 28 (3), 221–229.

Rockwood, K., Mogilner, A., Mitnitski, A., 2004. Changes with age in the distribution of a frailty index. Mech. Ageing Dev. 125 (7), 517–519.

Rockwood, K., Song, X., MacKnight, C., Bergman, H., Hogan, D.B., McDowell, I., Mitnitski, A., 2005. A global clinical measure of fitness and frailty in elderly people. Can. Med. Assoc. J. 173 (5), 489–495.

Roque, F.S., Jensen, P.B., Schmock, H., Dalgaard, M., Andreatta, M., Hansen, T., Søeby, K., Bredkjær, S., Juul, A., Werge, T., Jensen, L.J., Brunak, S., 2011. Using electronic patient records to discover disease correlations and stratify patient cohorts. PLOS Comput. Biol. 7 (8), 1–10.

Rubanova, Y., Chen, T.Q., Duvenaud, D., 2019. Latent Ordinary Differential Equations for Irregularly-Sampled Time Series. NeurIPS.

Rudin, C., 2019. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat. Mach. Intell. 1, 206–215.

Santana, P., 2000. Ageing in Portugal: regional iniquities in health and health care. Soc. Sci. Med. 50 (7–8), 1025–1036.

Schulam, P., Suchi, S., 2015. A framework for individualizing predictions of disease trajectories by exploiting multi-resolution structure. Adv. Neural. Inf. Process. Syst.

Schultz, M.B., Kane, A.E., Mitchell, S.J., MacArthur, M.R., Warner, E., Mitchell, J.R., Howlett, S.E., Bonkowski, M.S., Sinclair, D.A., 2019. Age and life expectancy clocks based on machine learning analysis of mouse frailty. bioRxiv 34 (6). https://doi.org/10.1101/2019.12.20.884452.

Seroude, L., Brummel, T., Kapahi, P., Benzer, S., 2002. Spatio-temporal analysis of gene expression during aging in Drosophila melanogaster. Aging Cell 1 (1), 47–56.

Steptoe, A., Breeze, E., Banks, J., Nazroo, J., 2014. Cohort profile: the English Longitudinal Study of Ageing. Int. J. Epidemiol. 42 (6), 1640–1648.

Stolz, E., Mayerl, H., Hoogendijk, E.O., Armstrong, J.J., Roller-Wirnsberger, R., Freidl, W., 2020. Acceleration of health deficit accumulation in late-life: Evidence of terminal decline in frailty index three years before death in the US Health and Retirement Study. medRxiv. https://doi.org/10.1101/2020.10.11.20210732.

Stroustrup, N., Anthony, W.E., Nash, Z.M., Gowda, V., Gomez, A., López-Moyado, I.F., Apfeld, J., Fontana, W., 2016. The temporal scaling of Caenorhabditis elegans ageing. Nature 530 (7588), 103–107.

Stubbings, G., Farrell, S., Mitnitski, A., Rockwood, K., Rutenberg, A., 2020. Informative frailty indices from binarized biomarkers. Biogerontology 70 (1), 1–11.

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., Liu, B., Matthews, P., Ong, G., Pell, J., Silman, A., Young, A., Sprosen, T., Peakman, T., Collins, R., 2015. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 12 (3), e1001779.

Sun, E.D., Michaels, T.C.T., Mahadevan, L., 2020. Optimal control of aging in complex networks. Proc. Natl. Acad. Sci. 117 (34), 20404–20410.

Swierczek, N.A., Giles, A.C., Rankin, C.H., Kerr, R.A., 2011. High-throughput behavioral analysis in C. elegans. Nat. Methods 8, 592–598.

Taneja, S., Mitnitski, A.B., Rockwood, K., Rutenberg, A.D., 2016. Dynamical network model for age-related health deficits and mortality. Phys. Rev. E. 93 (2), 022309–022311.

Tarkhov, A.E., Menshikov, L.I., Fedichev, P.O., 2017. Strehler-Mildvan correlation is a degenerate manifold of Gompertz fit. J. Theor. Biol. 416, 180–189.

Timonen, J., Mannerström, H., Vehtari, A., Lähdesmäki, H., 2020. lgpr: An Interpretable Nonparametric Method For Inferring Covariate Effects From Longitudinal Data arXiv 1912.03549.

Transtrum, M.K., Machta, B.B., Brown, K.S., Daniels, B.C., Myers, C.R., Sethna, J.P., 2015. Perspective: sloppiness and emergent theories in physics, biology, and beyond. J. Chem. Phys. 143 (1), 010901.

Vassy, J.L., Ho, Y.L., Honerlaw, J., Cho, K., Gaziano, J.M., Wilson, P.W.F., Gagnon, D.R., 2018. Yield and bias in defining a cohort study baseline from electronic health record data. J. Biomed. Inform. 78, 54–59.

Vural, D.C., Morrison, G., Mahadevan, L., 2014. Aging in complex interdependency networks. Phys Rev E 89, 022811.

Wallace, L.M.K., Theou, O., Pena, F., Rockwood, K., Andrew, M.K., 2015. Social vulnerability as a predictor of mortality and disability: cross-country differences in the survey of health, aging, and retirement in Europe (SHARE). Aging Clin. Exp. Res. 27 (3), 365–372.

Walsh, J.R., Smith, A.M., Pouliot, Y., Li-Bland, D., Loukianov, A., Fisher, C.K., 2020. Generating Digital Twins With Multiple Sclerosis Using Probabilistic Neural Networks arXiv 2002:02779v1.

Williams, D.R., 2005. The health of U.S. racial and ethnic populations. J. Gerontol. B-Psychol. 60 (Special Issue 2), S53–S62.

Xu, M., Pirtskhalava, T., Farr, J.N., Weigand, B.M., Palmer, A.K., Weivoda, M.M., Inman, C.L., Ogrodnik, M.B., Hachfeld, C.M., Fraser, D.G., Onken, J.L., Johnson, K.O., Verzosa, G.C., Langhi, L.G.P., Weigl, M., Giorgadze, N., LeBrasseur, N.K., Miller, J.D., Jurk, D., Singh, R.J., Allison, D.B., Ejima, K., Hubbard, G.B., Ikeno, Y., Cubro, H., Garovic, V.D., Hou, X., Weroha, S.J., Robbins, P.D., Niederhofer, L.J., Khosla, S., Tchkonia, T., Kirkland, J.L., 2018. Senolytics improve physical function and increase lifespan in old age. Nat. Med. 24 (8), 1246–1256.

Yashin, A.I., Iachine, I.A., Begun, A.S., 2000. Mortality modeling: a review. Math. Popul. Stud. 8 (4), 305–332.

Yashin, A.I., Arbeev, K.G., Akushevich, I., Kulminski, A., Akushevich, L., Ukraintseva, S.V., 2007. Stochastic model for analysis of longitudinal data on aging and mortality. Math. Biosci. 208, 538–551.

Yashin, A.I., Arbeev, K.G., Akushevich, I., Kulminski, A., Ukraintseva, S.V., Stallard, E., Land, K.C., 2012. The quadratic hazard model for analyzing longitudinal data on aging, health, and the life span. Phys. Life. Rev. 9, 177–188.

Yu, P., Song, X., Shi, J., Mitnitski, A., Tang, Z., Fang, X., Rockwood, K., 2012. Frailty and survival of older Chinese adults in urban and rural areas: results from the Beijing Longitudinal Study of Aging. Arch. Gerontol. Geriatr. 54, 3–8.

Zajacova, A., Dowd, J.B., 2011. Reliability of self-rated health in US adults. Am. J. Epidemiol. 174 (8), 977–983.

Zhang, W.B., Sinha, D.B., Pittman, W.E., Hvatum, E., Stroustrup, N., Pincus, Z., 2016. Extended twilight among isogenic C. elegans causes a disproportionate scaling between lifespan and health. Cell. Syst. 3 (4), 333-345.e4.

Zhavoronkov, A., Mamoshina, P., 2019. Deep aging clocks: the emergence of AI-based biomarkers of aging and longevity. Trends Pharmacol. Sci. 40 (8), 546–549.