Accelerating research on biological aging and mental health: Current challenges and future directions

Laura K.M. Han\textsuperscript{a,b}, Josine E. Verhoeven\textsuperscript{a}, Audrey R. Tyrka\textsuperscript{c}, Brenda W.J.H. Penninx\textsuperscript{a,b}, Owen M. Wolkowitz\textsuperscript{d}, Kristoffer N.T. Månsson\textsuperscript{e,g}, Daniel Lindqvist\textsuperscript{b,i}, Marco P. Boks\textsuperscript{k}, Dóra Révész\textsuperscript{l}, Synthia H. Mellon\textsuperscript{d}, Martin Picard\textsuperscript{m,n,o}\textsuperscript{⁎}

\textsuperscript{a} Amsterdam UMC, Vrije Universiteit Amsterdam, Department of Psychiatry, Amsterdam Public Health Research Institute, Oldenaller 1, the Netherlands
\textsuperscript{b} Amsterdam UMC, Vrije Universiteit Amsterdam, Department of Psychiatry, Amsterdam Neuroscience, De Boelelaan 1117, Amsterdam, the Netherlands
\textsuperscript{c} Butler Hospital and the Department of Psychiatry and Human Behavior, Warren Alpert Medical School of Brown University, Providence, RI, USA
\textsuperscript{d} Department of Psychiatry and Weill Institute for Neurosciences, University of California, San Francisco, School of Medicine, San Francisco, CA, USA
\textsuperscript{e} Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
\textsuperscript{f} Department of Psychology, Stockholm University, Stockholm, Sweden
\textsuperscript{g} Faculty of Medicine, Department of Clinical Sciences, Psychiatry, Lund University, Lund, Sweden
\textsuperscript{h} Department of Psychiatry, University of California San Francisco (UCSF) School of Medicine, San Francisco, CA, USA
\textsuperscript{i} Psychiatric Clinic, Lund, Division of Psychiatry, Lund, Sweden
\textsuperscript{j} Brain Center Rudolf Magnus, Department of Psychiatry, University Medical Center Utrecht, the Netherlands
\textsuperscript{k} Center of Research on Psychology in Somatic diseases (CoRPS), Department of Medical and Clinical Psychology, Tilburg University, Tilburg, the Netherlands
\textsuperscript{m} Department of Psychiatry, Division of Behavioral Medicine, Columbia University Medical Center, New York, NY, USA
\textsuperscript{n} Department of Neurology, H. Houson Merritt Center, Columbia Translational Neuroscience Initiative, Columbia University Medical Center, New York, NY, USA
\textsuperscript{o} Columbia Aging Center, Columbia University, New York, NY, USA

	extbf{ARTICLE INFO}

\textbf{Keywords:}
Biological age
Psychopathology
DNA methylation
Brain
Mitochondria
Telomere length

\textbf{ABSTRACT}

Aging is associated with complex biological changes that can be accelerated, slowed, or even temporarily reversed by biological and non-biological factors. This article focuses on the link between biological aging, psychological stressors, and mental illness. Rather than comprehensively reviewing this rapidly expanding field, we highlight challenges in this area of research and propose potential strategies to accelerate progress in this field. This effort requires the interaction of scientists across disciplines - including biology, psychiatry, psychology, and epidemiology; and across levels of analysis that emphasize different outcome measures - functional capacity, physiological, cellular, and molecular. Dialogues across disciplines and levels of analysis naturally lead to new opportunities for discovery but also to stimulating challenges. Some important challenges consist of 1) establishing the best objective and predictive biological age indicators or combinations of indicators, 2) identifying the basis for inter-individual differences in the rate of biological aging, and 3) examining to what extent interventions can delay, halt or temporarily reverse aging trajectories. Discovering how psychological states influence biological aging, and vice versa, has the potential to create novel and exciting opportunities for healthcare and possibly yield insights into the fundamental mechanisms that drive human aging.

1. Introduction

Aging is the strongest risk factor for many chronic illnesses, loss of functional capacity, and mortality (Fernandes et al., 2016). It is associated with complex biological changes, but there is no consensus on the very definition of aging, nor on the best methods to quantify it biologically (Xia et al., 2017). Chronological age is based on the passage of time and is invariable. But biological age may fall behind or else outpace,...
Currently, although we have some quantifiable measures of biological health and the aging process. Logical, and biological aging processes as a platform for the development and clinical domains. We see the intersection of experiential, psychological, and biological aging requiring dialogue across research and meaningful, valid, and reproducible measures of biological aging. Second, this interdisciplinary problem requires dialogue across research and clinical domains. We see the intersection of experiential, psychological, and biological aging processes as a platform for the development of new (and possibly radically different) concepts and measures that will more faithfully capture human health and the aging process. Currently, although we have some quantifiable measures of biological aging in humans - biological age indicators - we still know little about their causal role in the aging process, and about their modifiability by psychological states and psychopathology.

One important shared goal towards enhancing well-being across the lifespan is to understand aging as dynamic trajectories determined by a variety of factors. Some determinants of biological aging are pre-programmed (“intrinsic”; e.g., genetic), while others are affected by the environment (“extrinsic”; e.g., diet, adversity) (de Magalhães, 2012). Most definitions of biological aging include loss of function, increased propensity to certain diseases, and closer proximity to death (de Magalhães and Passos, 2018). Certain objective biological measures (or “clocks”) may also track biological aging. Development and validation of biological age indicators and clarification of their mediators and moderators are high priorities, since they may lead to a better understanding of the underpinnings of healthy and unhealthy aging trajectories. These indicators may also present proximal outcomes, or “early warning signs” that portend disease development and may provide a more sensitive platform to detect – and intervene upon – meaningful interactions between psychological, social, and bio-behavioral factors that influence aging trajectories and health outcomes.

Biological age indicators currently being investigated include telomere length (TL), epigenetic changes, alterations of mitochondrial function and mitochondrial DNA (mtDNA), age-related brain structure and function, and transcriptomic, metabonomic, and proteomic changes, among others (see, Cole et al., 2018; Jylhava et al., 2017; Xia et al., 2017 for recent reviews). Current topics of investigation include the nature of the inter-relationship of these biological age indicators, whether they measure the same or different aspects of biological aging, whether they are causally involved in the aging process, whether they have a causal role in disease and disorders, and the best ways to assess them. The possibility that the aging process is accelerated by chronic psychological stress and that it plays a role in the pathophysiology of some mental illnesses has been supported by observations that chronically stressed or psychiatrically ill individuals are at increased risk of acquiring specific age-related diseases and have a reduced life expectancy (Epel and Prather, 2018; Lindqvist et al., 2015; Penninx et al., 2013; Verhoeven et al., 2014a, b; Walker et al., 2015). But certain conceptual and methodological obstacles are impeding growth in this field and hinder replication of findings across laboratories.

Rather than comprehensively reviewing this rapidly expanding field, here we focus on highlighting various challenges and arising opportunities for this interdisciplinary endeavor. We conclude by proposing strategies to accelerate the progress of this field towards a predictive science that can enhance our understanding of the psychobiological factors that influence the aging process and lifespan.

2. The concept of biological aging: Definitions and obstacles

Chronic age is strictly quantitative and requires no more than a calendar to measure. Biological age is more elusive as it reflects the functional and biological condition of an individual. The difference between biological age and chronological age can indicate whether the individual’s biological state is “older” or “younger” than would be expected
for a given chronological age. This is often referred to as “accelerated” or “slowed” aging, respectively. However, cross-sectional assessments of biological age do not allow one to determine aging rates or to distinguish between “accelerated” and “premature” or “advanced” aging (Fig. 2). The rate of increase in biological aging over time may also be nonlinear, particularly in early life where the measured rate of aging may be more rapid than across adult life (Cole et al., 2018; Lohr et al., 2015).

Since it is not possible to directly assess the total biological state of a person, biological age indicators serve as proxies. Biological age indicators are functional, anatomical, biochemical, cellular or molecular measures that are correlated with age and that may reflect the health status of specific cell types and/or organ systems. The term biomarker, in contrast to indicators as defined here, is best used in the context of specific disease or health outcomes. By definition, biomarkers must exhibit both sensitivity and specificity in relation to the outcome we design them to predict (Abi-Dargham and Horga, 2016). For aging, a still broadly defined process compared to a disease that can be ascertained with certainty, the term indicator is rendered more appropriate. Although some biological age indicators have undoubtedly established their sensitivity to chronological age, few have convincingly demonstrated their specificity - that changes in their value occurs specifically in response to the aging process and not in response to other pathophysiological processes. Some biological age indicators are indeed modified by disease states independent of aging, and some as discussed below may notably be sensitive to psychological states, namely stress and psychopathology.

3. Biological age indicators

Aging is a multifaceted and complex process that manifests across multiple levels. In recent decades, measurements spanning each of these levels have been developed, reflecting our prevailing reductionist scientific approach to biomedical sciences. Here, rather than providing an exhaustive overview that can be found in recent reviews (Engelriet et al., 2013; Horvath and Raj, 2018; Jylhäva et al., 2017; López-Ortí et al., 2013; Wagner et al., 2016; Xia et al., 2017), we provide a selective overview of aging indicators commonly studied in relation to mental health (Table 1). This table, which is illustrative rather than comprehensive, also includes some emerging biological age indicators that reflect the development of omics technologies and of computational approaches to integrate multiple metrics into composite indices, as also discussed in Section 5. Due to space limitation, we do not cover self-reported age (Stephan et al., 2018), self-reported perception of aging (Levy et al., 2002), more specific brain measures (Cole, 2018), or inflammatory markers (Franceschi et al., 2000), which have also been associated with lifespan. We then discuss the practical limitations and conceptual challenges commonly associated with these measurements.

3.1. General limitations common to biological age indicators

There are limitations inherent to existing biological age indicators: they rely on specific organs or tissue types, can be confounded by cell type heterogeneity, and represent static measures of dynamic states (Fig. 3). These limitations apply to most molecular biological age indicators and should represent the foundation from which we design research projects, and interpret findings. However, these limitations are often not well understood and, instead, only considered post-hoc once data is collected and is being analyzed. Here, to be consistent with the logic whereby limitations inform research design, method development, and data interpretation, we discuss these challenges prior to the literature review. Recommendations to overcome some of these limitations are also discussed in Section 5.

3.1.1. Organs and tissue types

A common limitation of biological age indicators is that they are generally measured in one particular tissue and then used as a general age estimator for the person from whom the sample was obtained. However, it is unlikely that every tissue - peripheral and central - are entirely synchronized and that one tissue accurately reflects the biological age of all other tissues. Tissues from multiple organs have been found to age at different rates in terms of epigenetic age (Horvath et al., 2015). With age, mtDNA mutations also accumulate differently

![Cross-sectional data](https://example.com/cross-sectional-data.png)

**Cross-sectional data**
- Run linear regression across all individuals
- Determine residual for each person
- Value represents predicted biological age difference

**Accelerated Age** = residual > 0

**Decelerated Age** = residual < 0

![Longitudinal data](https://example.com/longitudinal-data.png)

**Longitudinal data**
- Measure biological age indicator at multiple time points (ideally ≥3) for the same individual
- Compute slope
- Difference with reference slope represents individualized rate of aging

**True aging acceleration** = Slope > 1

**True aging deceleration** = Slope < 1

**Fig. 2.** Computing age acceleration using cross-sectional and longitudinal data. (A) From cross-sectional data, accelerated aging is established when biological age is over-predicted relative to the chronological age reference (group regression line). (B) In longitudinal data, the rate of aging is directly determined from multiple measurements in the same person (also same tissue and cell type). The slope for each individual can be compared to the theoretical slope of 1 to ascertain true aging acceleration or deceleration.
<table>
<thead>
<tr>
<th>Category</th>
<th>Biological age indicator</th>
<th>Method</th>
<th>Evidence or mental health association</th>
</tr>
</thead>
<tbody>
<tr>
<td>Functional &amp; physiological measures</td>
<td>Walking speed/gait</td>
<td>Timed distance walked</td>
<td>Impaired in depression (García et al., 2002; Mikhailov et al., 2009); Inverse relationship with depression (C order et al., 2018; Liang et al., 2017)</td>
</tr>
<tr>
<td>Brain measures</td>
<td>Hand grip strength</td>
<td>Dynamometer</td>
<td>Inverse association with depression (Lee et al., 2018; Lever-van Milligen et al., 2017)</td>
</tr>
<tr>
<td>Brain measures</td>
<td>Lung function</td>
<td>Spirometer</td>
<td>Lower in depression (Lever-van Milligen et al., 2017)</td>
</tr>
<tr>
<td>Brain measures</td>
<td>Brain age</td>
<td>MRI</td>
<td>Higher in depression (Kiehl et al., 2009); MDD, BPD, and psychosis (Kolenic et al., 2018; Koutsouleris et al., 2014)</td>
</tr>
<tr>
<td>Brain measures</td>
<td>Brain age</td>
<td>MRI</td>
<td>Inconsistent in BD (Hajek et al., 2017; Nenadic et al., 2017)</td>
</tr>
<tr>
<td>Brain measures</td>
<td>Brain age</td>
<td>MRI</td>
<td>Higher with more negative life events (Hatton et al., 2018)</td>
</tr>
<tr>
<td>Cellular measures</td>
<td>Mitochondrial enzymatic activities</td>
<td>Respiratory capacity</td>
<td>Enzymatic activities correlated with previous day positive mood (Picard et al., 2018)</td>
</tr>
<tr>
<td>Cellular measures</td>
<td>Mitochondrial enzymatic activities</td>
<td>Respiratory capacity</td>
<td>Enzymatic activities: correlated with previous day positive mood (Picard et al., 2018)</td>
</tr>
<tr>
<td>Cellular measures</td>
<td>Mitochondrial enzymatic activities</td>
<td>Respiratory capacity</td>
<td>Enzymatic activities: correlated with previous day positive mood (Picard et al., 2018)</td>
</tr>
<tr>
<td>Molecular measures</td>
<td>Telomere length</td>
<td>Southern blot</td>
<td>Shorter in MDD (Schutte and Malouff, 2015), anxiety disorders (Malouff and Schutte, 2017), SCZ (Polho et al., 2015), PTSD (Li et al., 2017a,b)</td>
</tr>
<tr>
<td>Molecular measures</td>
<td>Telomere length</td>
<td>Quantitative fluorescent in situ hybridization Hybridization (Hannum et al., 2013; Horvath, 2013; Marioni et al., 2015)</td>
<td></td>
</tr>
<tr>
<td>Molecular measures</td>
<td>Epigenetic age</td>
<td>Microarray and sequencing</td>
<td>Increased in response to traumatic stress (Wolf et al., 2018), life events (Hatton et al., 2018)</td>
</tr>
<tr>
<td>Proteomic age</td>
<td>ccf-mtDNA qPCR</td>
<td>Positive</td>
<td>NC in plasma of suicidal patients (Lindqvist et al., 2016) and MDD (Lindqvist et al., 2018)</td>
</tr>
<tr>
<td>Metabolomic age</td>
<td>mtDNAcn qPCR</td>
<td>Inverse</td>
<td>Lower in plasma of suicidal patients (Lindqvist et al., 2016) and MDD (Lindqvist et al., 2018)</td>
</tr>
<tr>
<td>Metabolomic age</td>
<td>mtDNAcn qPCR</td>
<td>Inverse</td>
<td>Lower in plasma of suicidal patients (Lindqvist et al., 2016) and MDD (Lindqvist et al., 2018)</td>
</tr>
<tr>
<td>Metabolomic age</td>
<td>mtDNAcn qPCR</td>
<td>Inverse</td>
<td>Lower in plasma of suicidal patients (Lindqvist et al., 2016) and MDD (Lindqvist et al., 2018)</td>
</tr>
<tr>
<td>Metabolomic age</td>
<td>mtDNAcn qPCR</td>
<td>Inverse</td>
<td>Lower in plasma of suicidal patients (Lindqvist et al., 2016) and MDD (Lindqvist et al., 2018)</td>
</tr>
<tr>
<td>Metabolomic age</td>
<td>mtDNAcn qPCR</td>
<td>Inverse</td>
<td>Lower in plasma of suicidal patients (Lindqvist et al., 2016) and MDD (Lindqvist et al., 2018)</td>
</tr>
<tr>
<td>Metabolomic age</td>
<td>mtDNAcn qPCR</td>
<td>Inverse</td>
<td>Lower in plasma of suicidal patients (Lindqvist et al., 2016) and MDD (Lindqvist et al., 2018)</td>
</tr>
<tr>
<td>Metabolomic age</td>
<td>mtDNAcn qPCR</td>
<td>Inverse</td>
<td>Lower in plasma of suicidal patients (Lindqvist et al., 2016) and MDD (Lindqvist et al., 2018)</td>
</tr>
<tr>
<td>Metabolomic age</td>
<td>mtDNAcn qPCR</td>
<td>Inverse</td>
<td>Lower in plasma of suicidal patients (Lindqvist et al., 2016) and MDD (Lindqvist et al., 2018)</td>
</tr>
<tr>
<td>Metabolomic age</td>
<td>mtDNAcn qPCR</td>
<td>Inverse</td>
<td>Lower in plasma of suicidal patients (Lindqvist et al., 2016) and MDD (Lindqvist et al., 2018)</td>
</tr>
</tbody>
</table>
between brain regions (Bender et al., 2006; Corral-Debrinski et al., 1992; Picard and Hirano, 2016) and even between cells of a given organ (Vincent et al., 2018). Within the brain, TL also varies between different cortical areas (Mamdani et al., 2015). This limitation may also apply to measures of functional capacity. For example, there is relatively poor agreement between muscle strength measured from handgrip or knee extension, suggesting that the most commonly used metric of muscle strength - handgrip strength - is not a proxy for overall muscle strength (Yeung et al., 2018). Developing approaches to measure and perhaps capitalize on the heterogeneous nature of aging dynamics across tissues requires further research.

### 3.1.2. Cellular composition and heterogeneity

Tissues such as the brain, heart, muscles, and the blood are composed of multiple different cell types. Although all cell types have the same genome, they show unique epigenetic, morphological, functional, and molecular differences relevant to biological age indicators. For example, different blood cell types have different epigenomes (Reinius et al., 2012), telomerase activity and TL (Boeck et al., 2018b, 2018d; Lin et al., 2010), and mitochondrial respiratory capacity (Chacko et al., 2013). These intrinsic differences are inevitable and may introduce bias and confound findings when assessed individually. Similar limitations of cellular heterogeneity also apply to saliva, skin, and any tissue such as placenta, brain, and others. However, in many cases, the relative cellular composition of these tissues is poorly characterized or methods may not be available to effectively disentangle cellular composition effects, relative to the true biological aging signal.

### 3.1.3. Static indicators of dynamic processes

Most biological age indicators reflect the current state of the organism, and the biological age of the sampled tissue and cells, at the moment of collection. In many cases, it is unknown how dynamic these markers are. In other words, how much they change from day-to-day, across the day (i.e., diurnal variation), or sometimes even within minutes, as is the case for neuroendocrine mediators and blood-based metabolites. It is generally assumed that most biological age indicators (and the specific measures that compose some of them) are largely stable, changing slowly over the course of years, but that assumption has gone untested or proven false for most indicators listed in Table 1. Unrecognized, unmeasured, or uncontrolled variability of biological age indicators due to regular or irregular changes over time has two undesirable effects: it introduces noise that cannot be accounted for, and possibly limits our interpretation of the downstream result. Biological age indicators may follow different trajectories over time and may theoretically be differentially sensitive to behaviors such as sleep, exercise, diet, meditation, and others. Studies with frequently assessed repeated measures of biological age indicators (e.g. over hours, days, months, and years) will be necessary to establish the temporal kinetics for existing and new biological age indicators.

### 3.2. Specific limitations to measurements of biological aging

#### 3.2.1. Limitations of functional capacity and physiological measures

A major advantage of functional capacity and physiological measures is the high efficiency in terms of costs and collection as well as their integrative informativeness, specifically compared to blood-based, molecular and DNA-based measurements. However, the predictive value of objective functional and physiological measures on mortality has mostly been reported in older populations (Cooper et al., 2010) and in middle-aged populations (Rantanen et al., 2000). Except for some data reporting associations between handgrip strength and mortality in male adolescents (Ortega et al., 2012), whether functional capacity measures such as walking speed are sensitive to aging and predictive of morbidity and mortality in younger populations largely remains to be established.

#### 3.2.2. Limitations of brain function- and structure-based measures

Normal aging is accompanied by brain atrophy and loss of brain tissue volume, which can be quantified non-invasively with magnetic
resonance imaging (MRI). Voxel-based morphometry and surface-based analysis are two commonly used image preprocessing techniques, which may yield divergent results (Chung et al., 2017; Clarkson et al., 2011). Moreover, the macroscopic volumetric changes observed in T1- and T2-weighted MR imaging reflect microscopic changes at the tissue and cellular levels, and in many circumstances possibly represent an aggregate of multiple cellular mechanisms related to synapses, neurons, and glial cells (Tardif et al., 2016). Thus, what changes in brain volume represent is not fully understood.

Age-related differences in brain function can also be detected with functional connectivity (e.g., Geerligs et al., 2015) and novel analytics on brain response are also available. For example, Garrett et al. (2010) showed that the age-predictive power of the brain’s signal variability was five times higher than that of the conventional method of assessing the average signal across time. But functional MRI data has low signal-to-noise ratio, and movement artifacts are one source of such noise. Because individuals of different ages might move differently in response to assessments, movement artifacts are a possible confounder in many study designs. Implementing methods to systematically review individual participant’s images and manually separate noise from signal (Griffanti et al., 2017) could be a useful technique to minimize artifacts. There is evidence that manually cleaned BOLD-fMRI data, compared to data preprocessed with conventional automatic methods, better predicts chronological age (Garrett et al., 2010), emphasizing the importance of data quality and pre-processing procedures in conclusions derived from brain-based age indicators.

A popular approach in neuroscience is to use statistical approaches to translate complex whole-brain multivariate patterns of aging into a single outcome (Luders et al., 2016), the so-called “brain age” (see Cole and Franke, 2017 for details). Brain age algorithms (Ayache et al., 2018; Ball et al., 2017; Cole et al., 2016; Gaser et al., 2013; Liem, 2016; Schnack et al., 2016) generate accurate individual age predictions in healthy controls, but show greater prediction errors when applied to patient groups (Cole et al., 2018). Within this framework, neuropathology may be reflected by the trajectory of aberrant normal aging, rather than a different deteriorating pattern of pathology. Gutierrez Becker et al. (2018) show that Gaussian Process uncertainty in age estimation may yield a better separation between cases and healthy individuals than the prediction error. Nevertheless, brain age models have high reliability in terms of test-retest performance at both same and different scanners (Cole et al., 2016; Franke and Gaser, 2012), and have shown biologically meaningful associations with health, clinical, and neuropsychiatric phenotypes (Cole et al., 2018).

3.2.3. Limitations of cellular measures

Although cellular measures of aging have been used widely in the laboratory setting, they are seldom applied to human (clinical, epide-miological) research. For instance, replicative senescence, a cellular measure of aging, involves monitoring cells grown in culture, and counting the number of cells over time (e.g., Lawless et al., 2010). This enables the investigator to count the number of times cells divide (i.e., total population doublings or “hayflick limit”), and determine the time required per cell division (i.e., population doubling time), which increases as cells age and divide more slowly. It should be noted that although senescence - defined as the loss of the ability of a cell to grow or divide - is associated with aging, it is not equivalent and can be dissociated from chronological age. Indeed, other factors such as irradiation can specifically induce senescence, even in chronologically young cells, thus reflecting biological aging.

The assessment of cellular bioenergetics, particularly mitochondrial content and functions, represents another domain of cellular aging measures. These indicators reflect the ability of cells to generate energy through oxygen-dependent mechanisms (for a review, see Picard and McEwen, 2018a). Respiration can be measured in whole cells (Boeck et al., 2018a) where it mostly reflects cellular energy demand, or in permeabilized cells (Ehinger et al., 2015) and isolated mitochondria where the intrinsic function of the organelle can be directly assessed independent of cellular contributions (Tyrrell et al., 2015). A major limitation of functional measures on intact cells and mitochondria is that measurements must be performed rapidly after blood sampling (within minutes to hours), which limits throughput and increases technical variability between samples, requiring exceptional standardization of procedures.

Other approaches relying on lysates (homogenized cells or mitochondria) from frozen samples allow measurement of enzymatic activity (e.g., telomerase, mitochondrial respiratory chain complexes) for multiple samples at once or in large batches (Picard et al., 2018). A ubiquitous limitation to all measures of biological activity (as opposed to inert molecules) is the degradation of cellular and enzymatic activities over time when samples are stored under suboptimal conditions. An important unknown in this field is the degree to which storage conditions, and especially the length of time a sample resides in a freezer, contributes to changes in assay results. This issue is worthy of applied investigation, since there is commonly a trade-off between freezing samples for shorter periods of time vs. freezing samples for longer periods of time to minimize inter-assay variability in assaying sequential frozen batches. In contrast to molecular analytes that are mostly or fully preserved at −80°C, samples destined for functional measurements should be stored in liquid nitrogen (< −150°C).

3.2.4. Limitations of molecular measures

A large fraction of the most widely used biological age indicators is molecular in nature. They include DNA methylation (DNAm), metabolites, proteins, TL, mtDNAcn, circulating cell-free mtDNA (ccf-mtDNA), mtDNA damage, and others. One important consideration to all molecular measures is that inadequate handling of fresh samples can alter the concentration of various analytes, particularly metabolites. For example, whereas DNA markers are believed to be quite stable over minutes to days, blood glucose concentration decreases within minutes when the blood is left at room temperature (Chan et al., 1989), owing to metabolic activities of white and red blood cells. The same must also apply to other metabolites that are detected by metabolomics. Gene expression assessed from messenger RNA transcript levels is also subject to rapid degradation and special care must be applied to blood destined to transcriptomic analyses (Kågedal et al., 2005). These effects are minimized by rapid separation of the liquid and cellular components of whole blood by centrifugation immediately after blood draw, refrigeration (immediate storage of samples on wet ice, 4°C), and subsequently freezing biological samples in a timely fashion.

Below we discuss specific molecular biological age indicators that have been subject of considerable research in relation to stress and psychopathology. Although exciting new findings from proteomic (Menni et al., 2015; Tanaka et al., 2018) and metabolomic (Chaleckis et al., 2016; Hertel et al., 2016) signatures of aging are beginning to arise, they have not been examined in relation to psychological factors. In this section, we focus our discussion on DNAm, TL, and mtDNAcn.

3.2.4.1. DNA methylation and epigenetic age. To date, several epigenetic age estimators have been developed from e.g. whole blood (Hannum et al., 2013; Levine et al., 2018; Weidner et al., 2014), neonatal cord blood and blood spots (Knight et al., 2017), and skin and blood cells (Horvath et al., 2018). Many potential confounders may cause technical variation in DNAm studies, of which population stratification and genetic ancestry are major contributors (Heijmans and Mill, 2012; Ratanatharthorn et al., 2017; Susser et al., 2016). Therefore, there is also reason to assume that genetic variation impacts epigenetic age estimates, particularly considering recent studies that report strong genetic links (Lu et al., 2018). Other potential confounders include smoking (Elliott et al., 2014), sex, and prenatal factors (Simpkin et al., 2016). Technically, DNAm arrays can show large variations between individual arrays and batches, and methods have been designed to statistically correct for these prior to analyses (Akulenko et al., 2016; Price and Robinson, 2018).
While the validity of different methylation-based predictors is questioned, applications of the original pan-tissue Horvath clock (Horvath, 2013) have been successful across tissue types and proven accurate even in embryonic brain samples (Spiers et al., 2015). The latest developed skin and blood predictor seems to be even more robust across tissue types (Horvath et al., 2018). More recently, methods using as little as 3–10 CpG sites from blood samples also accurately predict age (Li et al., 2018; Weidner et al., 2014) and mortality risk scores (Gao et al., 2018).

Whether statistical adjustment for cell type composition should be uniformly applied to whole blood-derived DNA to achieve optimal age prediction is the subject of ongoing debate. Whereas it has been argued that the Horvath method incorporates the estimation of cell type composition from blood and may not require cell type adjustment, some have shown that intrinsic (without adjustment for cell type composition) and extrinsic (with adjustment for cell type composition) age estimates may have a different biological meaning (Chen et al., 2017). Also important to mention here is that other DNAm-based indicators of aging have been developed that are trained on phenotypic markers of age (DNAm PhenoAge) rather than chronological age, leading to improved predicted risk of mortality (Levine et al., 2018). Briefly, phenotypic age is a combination of chronological age and nine disease-related biomarkers selected based on their association with mortality. The sensitivity of epigenetic predictors to psychosocial stress and psychopathology remains a gap of knowledge.

3.2.4.2. Telomeres length. Issues related to TL assessments have been presented elsewhere (Aubert et al., 2012) and will not be discussed in detail here. In general, we will note that multiple different assays exist, which vary in their cost, required volume, applicability on frozen samples, and throughput. Technically, these are important considerations that impact the feasibility of clinical and epidemiological studies. Moreover, because of the relatively inexpensive and high-throughput capacity of qPCR-based methods, TL has frequently been measured on total DNA extracted from cell mixtures, which can be derived from a variety of sources such as buccal swabs (which include both epithelial cells and leukocytes) and whole blood (which include a variety of leukocytes). For these measurements, and as for mtDNA measurements below, the limitations presented in Section 3.1 are particularly important.

3.2.4.3. mtDNA copy number and circulating cell-free mtDNA. Counting the number of mtDNA molecules per cell, or mtDNAcn, can indirectly provide an indication of the bioenergetic state of the cell. The mtDNAcn measurements are based on either qPCR (e.g., Tyrka et al., 2016) or derived from whole exome or genome sequencing data (e.g., Cai et al., 2015) where "counts" of both the mitochondrial and nuclear genomes are estimated. The ratio of mtDNA and nuclear DNA (nDNA) is then multiplied by 2 to account for the diploid nature of the nuclear genome and taken as mtDNAcn (Malik et al., 2011). Here, given that cells with different metabolic demand can differ by as much as an order of magnitude in their content of mtDNA, cell type differences may have a particularly profound effect on this measure. When applied to a homogenous cell population, mtDNAcn can provide valuable information. However, most reported studies with mtDNAcn have relied on whole blood DNA, which is confounded by cell type heterogeneity, and by the presence of platelets. Platelets do not have nDNA, but have mtDNA, which artificially inflates mtDNA copy number in whole blood preparations (Hurtado-Roca et al., 2016; Urata et al., 2008). In tissues with less heterogeneity than blood, such as skeletal muscle, some have observed no difference in mtDNAcn between young and old individuals (Miller et al., 2003). The rate of decline in mtDNAcn per year also varies widely between studies, possibly as a result of differences in methodology and tissue source.

Similarly, measures of ccf-mtDNA are sensitive to cellular contamination and biological sample used. Serum (post-coagulation fraction of whole blood) may contain substantially more ccf-mtDNA than plasma (liquid fraction collected with an anticoagulant) (Xia et al., 2009). Sufficient centrifugation speed and time are required to successfully eliminate cells, particularly platelets, that could artificially inflate serum or plasma ccf-mtDNA (Nakahira et al., 2013). In studies where blood samples were not centrifuged at sufficient speeds, ccf-mtDNA levels are reportedly higher, making platelet contamination the most likely contributor to measured mtDNA levels and thus complicating interpretation of these results.

3.3. Associations among biological age indicators

While there is preliminary evidence for some cross-correlations among the different biological aging indicators, few have been examined in relation to other indicators. This highlights the need for examining multiple markers in an integrative study, as e.g. Belsky et al. (2018) recently showed low agreement between eleven quantifications of biological aging, with only modest associations to e.g. physical functioning and cognitive decline. Other studies suggest that TL is correlated to mtDNAcn (Tyrka et al., 2015), but the direction of stress and psychopathology effects with mtDNAcn or mitochondrial content (citrate synthase) and TL may vary (Boeck et al., 2018d; Cai et al., 2015; Picard et al., 2018; Tyrka et al., 2016). TL is not correlated with epigenetic age (Breitling et al., 2016; Han et al., 2018; Marioni et al., 2016), although cell type composition adjustments may reveal a modest association (Chen et al., 2017). Both epigenetic age and TL seem uncorrelated to brain age, and no associations were found between brain predicted age difference (brain-PAD = predicted brain age - chronological age) and epigenetic predicted age difference (Cole et al., 2017). The correlations between the Hamann and Horvath clocks vary from relatively strong ($r = 0.76$) to moderate ($r = 0.37$) in independent studies (Belsky et al., 2018; Chen et al., 2016), and both clocks showed modest correlations ($0.10 \sim 0.33$) to the transcriptomic age indicator by Peters et al. (2015). The microRNA age indicator of Huan et al. (2018) was modestly correlated to epigenetic age ($r = 0.3$) and microRNA expression ($r = 0.2$). Cross-correlations between metabolic/proteomic aging and other biological aging indicators remain to be explored.

4. Do psychological stress and psychopathology influence biological aging?

Nearly 15 years have elapsed since Epel and colleagues first described the association of psychosocial stress with short leukocyte telomeres in a sample of healthy premenopausal mothers of a chronically ill child and mothers of healthy children (Epel et al., 2004). The association of shortened telomeres with stress exposure has since been replicated in a wide variety of studies, and this observation stimulated several related lines of research examining the relationship between various forms of stress exposure or perceived stress and TL across the lifespan. Several excellent qualitative reviews describe and critically review this literature (Entringer et al., 2018; Epel and Prather, 2018; Price et al., 2013; Shalev et al., 2013; Shields and Slavich, 2017) and meta-analyses now quantify the magnitude of these associations and identify potential moderators of effects (e.g., Hansen et al., 2017; Mathur et al., 2016; Ridout et al., 2018, 2016; Schutte and Malouff, 2015, 2014). More recently, an appreciation of the role of mitochondria in the acute stress response and chronic allostatic load (Picard et al., 2014; Picard and McEwen, 2018a, 2018b) has led to investigations of the association of psychological states, stress exposure in relation to mitochondrial functions and mtDNA (Boeck et al., 2016; Cai et al., 2015; Picard et al., 2018; Tyrka et al., 2016). Rather than being a comprehensive review, this section provides a brief overview of this field, emphasizing recent developments.

4.1. What is (psychological) stress?

Broadly defined, “stress” is the condition of being subjected to a stimulus (i.e., stressor) that invokes a response requiring the use of
resources to adapt or cope (Monroe, 2008; Shields and Slavich, 2017). “Stress” may refer to particular life events (e.g., job loss, death of a loved one, assault), contexts that are experienced as stressful or contain numerous stressors (e.g., poverty, neighborhood violence, famine) or the psychological or biological response to such an event or exposure (i.e., stress response, perceived stress) (Epel et al., 2018). Characteristics of the stressor(s) and the stress response may be important determinants of the biological response or adaptation and account for heterogeneity in the literature on stress and aging.

While some stressors occur in isolation, it is important to recognize that stress-inducing contexts, exposures, and perceptions of stress often covary, for example when families living in poverty experience neighborhood violence and feel unsafe. In addition, the level of perceived stress may vary substantially within a group exposed to the same stressor. Determinants of psychological and biological stress responses include the nature of the stressor in terms of type, scope, severity, chronicity, and how predictable and controllable the stressor is. Individual and social characteristics influencing the level of perceived stress include social, financial, cognitive, emotional, and behavioral resources for coping with, controlling, avoiding, and compensating for stressors (Epel et al., 2018).

4.2. Stress and biological aging: Evidence for a stress-aging axis involving telomeres and mitochondria

The literature on the association of stressors and perceived stress on TL is now sufficiently large that a number of meta-analyses have been conducted on the topic. Meta-analyses of the association between TL and childhood psychosocial stressors document significant effects that vary from small to medium in size (Hanssen et al., 2017; Li et al., 2017a; b; Ridout et al., 2018). Moderator analyses suggest larger effects for studies that examine more severe exposures (Hanssen et al., 2017) and those that include a wide range of adversity types (Ridout et al., 2018). In addition to cross-sectional investigations, a longitudinal study (Shalev et al., 2012) found that exposure to violence over a 5-year period in childhood predicted greater TL attrition, suggesting the possibility of a causal relationship. The biological mechanisms whereby adverse or positive experiences exert their lasting health effects remain mostly unknown. Effects on the germline and stem cells reserves, metabolic reprogramming, and rewiring of neural networks and brain circuitry are among many areas that deserve further research.

Turning to stressors that occur in adulthood, a significant association between perceived stress and shorter TL has been documented in meta-analyses, though this effect ranged from very small to medium in size (Hanssen et al., 2017; Sim et al., 2016) in size. Several studies have also shown associations of shorter TL with measures of severe or cumulative stress exposure in adulthood (for reviews see Shalev et al., 2013; Epel and Prather, 2018; Oliveira et al., 2016). Although some studies suggest that childhood adversity may account for associations between TL and adult stressors (Puterman et al., 2016; Révézés et al., 2016), there is also evidence that stressors experienced in adulthood prospectively predict telomere attrition (Puterman et al., 2015; van Ockenburg et al., 2015).

Given the growing literature demonstrating the central role that mitochondria play in the stress response and the aging process, recent studies have examined the association of early life stress with measures of mitochondrial function or mtDNAcn. Although mtDNAcn is not a measure of mitochondrial function and is impossible to interpret on its own, it is easily measured from stored DNA and has been measured in different studies. For example, childhood trauma or adversity, as well as adult psychopathology have been linked to higher mtDNAcn (Cai et al., 2015; Tyrla et al., 2016). In a small study of postpartum women, early life adversity was associated with greater cellular respiration reflecting increased cellular energy demand, which in turn was positively correlated with levels of pro-inflammatory cytokines and childhood maltreatment (Boeck et al., 2016). In a study of caregiving stress, caregivers were found to have reductions in a functional index of mitochondrial health (MHI) in blood leukocytes. Mitochondrial health was operationalized as a multivariate index designed to reflect functional capacity on a “per mitochondrial” basis. In this first study of MHI in mixed human leukocytes, the index included biochemical enzymatic activities for three mitochondrial enzymes and mtDNAcn. Using this composite index as an outcome, this study found that positive mood was associated with higher MHI and was a mediator of the association between caregiving and MHI (Picard et al., 2018). Another study found that suicide attempters have significantly higher plasma levels of ccf-mtDNA (Lindqvist et al., 2016) and another study found elevated ccf-mtDNA levels in individuals with major depressive disorder (MDD) (Lindqvist et al., 2018).

A limitation of this body of research is that certain behavioral, psychiatric and medical conditions frequently co-occur with stress exposures and covary with TL and other biological processes central to aging, and thus may have confounding effects. These include smoking, obesity, dietary influences, anxiety, depressed mood, post-traumatic stress disorder (PTSD), medications, and cardiometabolic conditions, among others. These influences are not uniformly assessed, excluded, or statistically controlled. A meta-analysis on the association of early adversity and TL identified that the magnitude of the effect was smaller in studies that included participants with medical or psychiatric conditions, and participants on medications (Ridout et al., 2018). This finding suggests that the relationship between these conditions and shortened telomeres might obscure the effect of stress exposure, or, alternatively, that the psychiatric and medical conditions may be primarily responsible for some of the telomere effect (Epel and Prather, 2018). Thus, the complex inter-relationships among these exposures, behavioral factors, and health conditions should be carefully considered when designing studies and analyzing and interpreting results.

4.3. Psychopathology and biological aging

Psychiatric disorders are associated with increased risk of aging-related medical conditions, including cardiovascular disease, stroke, dementia, diabetes, and obesity (Penninx et al., 2013; Viron and Stern, 2010), and early mortality (Walker et al., 2015). While part of the association may be explained by differences in health behaviors, because individuals with psychiatric disorders are more likely to smoke, drink alcohol, eat poorly, and exercise less than others (van Gool et al., 2007), associations between psychiatric disorder status and medical morbidity remain significant after adjusting for these factors. This has led to the hypothesis that psychiatric conditions may induce or result from accelerated or premature biological aging. As reviewed here, there are multiple biological age indicators, including a range of cellular and molecular measures such as TL, mitochondrial dysfunction, oxidative stress, gene expression, and others (López-Otín et al., 2013). Notably, inflammation is also a widely-used indicator of biological aging (Baylis et al., 2013), coined as “inflammaging” by Franceschi et al. (2000). However, because of its elaborate discussion elsewhere (Franceschi et al., 2018b; Fulop et al., 2018), as well as in respect to mental health (Diniz and Vieira, 2018), we do not discuss it in great detail here. Overall, the following section will review the most frequently studied biological age indicators in epidemiological psychiatric research, including TL, epigenetic age, brain age, and to a lesser extent pro-inflammatory cytokines.

4.3.1. Associations of psychiatric disorders and biological age indicators

Simon and colleagues (Simon et al., 2006) were the first to report a relationship between psychiatric disorders and shorter telomeres in a sample that included MDD, bipolar disorder (BD) and anxiety disorder patients. Since then, a large number of studies have been conducted in an assortment of psychiatric disorders. MDD is among the most frequently studied disorders in this context, possibly as a consequence of its relatively well-documented associations with dysregulated physical
health (Penninx et al., 2013). Several meta-analyses, the largest one containing > 34,000 subjects from 38 studies, summarized the results and provided consistent evidence of an inverse association between TL and depression, generally with small to medium effect sizes (Darrow et al., 2016; Ridout et al., 2016; Schutte and Malouff, 2015). Similar meta-analytic results were found for anxiety disorders (N > 19,000) (Malouff and Schutte, 2017), and PTSD (N > 3800) (Li et al., 2017a, b). BD, schizophrenia and other psychotic disorders have been less extensively examined. A meta-analysis including 1100 subjects from 7 studies found no difference in TL between BD cases and controls (Colpo et al., 2015). Two meta-analyses on schizophrenia of 1200 and 1600 subjects, respectively, found small effects for TL differences (Polho et al., 2015; Rao et al., 2016).

The epigenetic age indicator is most frequently examined in individuals with PTSD. A meta-analysis using data from 9 cohorts (combined N = 2186) found significant, albeit small, associations of greater epigenetic age with traumatic stress, but not with PTSD diagnosis (Wolf et al., 2018). Other studies also found such relations between PTSD and epigenetic aging using the Horvath predictor (Boks et al., 2015; Mehta et al., 2018; Zannas et al., 2015). Because PTSD is associated with alterations in the HPA axis and glucocorticoid signaling, these findings are consistent with enrichment of CpG in glucocorticoid response elements in the Horvath pan tissue epigenetic age clock (Zannas et al., 2015). Moreover, two recent studies for the first time showed “older” epigenetic age in MDD patients versus controls (Han et al., 2018; Whalley et al., 2017), while this is not seen in schizophrenia (McKinney et al., 2017b; Voisey et al., 2017).

Chronic, low-grade inflammation that increases on average with age is captured in the term “inflammaging” - a pro-inflammatory state proposed to contribute to the pathogenesis of age-related diseases (Franceschi et al., 2018a). An increase in the inflammatory response, together with microglial activation, in turn, can contribute to psychiatric diseases, such as MDD, schizophrenia, BD, and autism (Khandaker et al., 2015; Martínez-Ceogoitabengoa et al., 2016; Réus et al., 2015). The mechanisms responsible for increased inflammation in mood disorders remains poorly understood. Recent evidence suggests that ccf-mtDNA could contribute to this pro-inflammatory state, although evidence is mixed (Kageyama et al., 2018; Lindqvist et al., 2018, 2016). Related to the bacterial origin of mitochondria, the mtDNA is immunogenic. Released mtDNA molecules thereby act as damage-associated molecular patterns (DAMPs) recognized by toll-like receptors on immune cells and trigger immune cell activation (West and Shadel, 2016). Inflammaging could in part be due to increased ccf-mtDNA in older individuals (Pinti et al., 2014).

Increased plasma levels of ccf-mtDNA have been reported in suicidal and depressed patients (Lindqvist et al., 2016). Worse response to an antidepressant was associated with increasing ccf-mtDNA levels during the treatment course, and ccf-mtDNA was correlated with antioxidant enzyme glutathione peroxidase, possibly as a result of a compensatory response to cellular oxidative stress (Lindqvist et al., 2018). An experimental study using psychological stress induction in healthy middle-aged individuals also demonstrated that an acute bout of psychological stress may be sufficient to elicit a 2–3 fold increase in serum ccf-mtDNA within 30 min, suggesting that ccf-mtDNA is dynamically regulated (Trumpf et al., 2019). Consistent with previous findings linking ccf-mtDNA levels to cortisol levels following a dexamethasone suppression test (Lindqvist et al., 2016), glucocorticoid stimulation of human cells (fibroblasts) induced the release of mtDNA by mitochondria within minutes (Trumpf et al., 2019). Thus, the causes of elevated ccf-mtDNA in certain psychiatric conditions remain unknown, although clinical and cellular studies suggest that canonical neuroendocrine stress mediators, including but not limited to glucocorticoids, may be implicated.

In addition, associations have been reported between “older” brain age and several psychiatric disorders (for an overview, see e.g. Cole et al., 2018) such as borderline personality disorder (Nenadic et al., 2017), schizophrenia (Koutsouleris et al., 2014; Nenadic et al., 2017; Schnack et al., 2016), and first-episode psychosis (Kolenic et al., 2018), but findings were less consistent in BD. One relatively small study by Koutsouleris et al. (2014) showed a higher brain-PAD of + 4.0 years in MDD (N = 104). However, a preliminary study by Kaufmann et al. (2018) finds increased brain age in schizophrenia (Cohen’s d = 0.55) and BD (d = 0.30), but not MDD (N = 211, d = 0.10). In addition, these authors suggest that the brain age gap is a genetically modulated trait that is heritable and overlaps with polygenic architecture observed in common brain disorders. Another preliminary study from the Enhancing NeuroImaging Genetics through Meta-Analysis (ENIGMA) MDD working group showed a small but significant brain-PAD in MDD (d = 0.12) in a large mega-analysis including > 8,000 cases and controls (Han et al., 2019). There is also preliminary evidence of an association between psychiatric pathology and glycomic-based biological age indicators, which represent sugar-based modifications of proteins, RNA, and DNA molecules. Two studies have shown altered protein N-glycosylation profiles in female patients with MDD (Boeck et al., 2018c) and in PTSD (Moreno-Villanueva et al., 2013), indicative of advanced aging at the glycomic level. Associations with other omics-based indicators (transcriptomics, proteomics, metabolomics) and their interrelation remain to be explored.

4.3.2. Biological aging and psychopathology: the chicken or the egg?

While robust cross-sectional associations between psychiatric disorders and biological aging have been documented - at least for TL - the nature and direction of these associations remain unclear. Longitudinal studies have found mixed effects (see, e.g. Boks et al., 2015; Maniates et al., 2018; Shalev et al., 2014; Vance et al., 2018; Verhoeven et al., 2016). It is currently unknown whether 1) psychopathology-associated physiological disturbances accelerate biological aging, 2) premature biological aging antedates and is a vulnerability factor that causes psychopathology, or alternatively, 3) psychopathology and biological aging processes share underlying etiological roots, such as shared genetic risks, and happen to be correlated without a causal link between them. Recent studies using genomic and causal inference tools are being developed to overcome this limitation of observational studies (e.g., Wiitm-Andersen et al., 2017). This challenge, among others, as well as recommendations to move the field towards a predictive science, are discussed in details in Section 6.

5. Clinical implications

5.1. Disorder-specific or transdiagnostic phenomenon?

In a large meta-analysis considering multiple psychiatric disorders and TL, including depressive and anxiety disorders, PTSD, bipolar and psychotic disorder, no difference in effect sizes between disorders was found (Darrow et al., 2016). This indicates that different DSM-based diagnoses may not be associated with meaningful differences in biological aging (Lindqvist et al., 2015). Furthermore, several studies showed that short TL is associated with the same physiological dysregulations that are found in some but not all persons with psychiatric disorders. These include increased inflammation, oxidative stress markers, dysregulated HPA-axis, and metabolic dysregulations (Lee et al., 2011; Monickaraj et al., 2012; Wikgren et al., 2012; Zhang et al., 2016). While the degree to which telomeres are causally related to these mechanisms is unknown, this is suggestive of pathways through which telomere shortening and psychiatric disorders are interrelated, that is not limited to one diagnostic category. The current evidence suggests that short TL may be a non-specific biological marker for conditions in which people experience chronic psychological or physiological stress, rather than being a marker of a specific psychiatric condition. Similarly, the downstream biological ramifications of different disorders also overlap, with alterations at the organs and systems level (e.g. brain aging patterns, Cole, 2018). This evidence leaves us to consider that these indicators are not disease-specific nor suitable as diagnostic tools, but rather general indicators of psychopathology or abnormal mental states.
5.2. Biological age indicators as predictors of treatment outcome

Only a small number of studies have investigated whether biological aging indicators predict antidepressant treatment response. The first study suggesting such a link showed in a small sample of previously unmedicated MDD subjects that low baseline telomerase activity, and a greater increase in telomerase activity during eight weeks of selective serotonin reuptake inhibitors (SSRI) treatment were associated to superior clinical outcome (Wolkowitz et al., 2012). However, this study lacked a control condition, leaving open the possibility that naturally different clinical trajectories contributed to these effects. Nevertheless, these findings suggested that depressed patients with relatively low baseline telomerase activity may most benefit from therapies that may secondarily induce telomerase activation, and that telomerase activation may represent a mechanism of antidepressant action, consistent with several animal studies (reviewed in Bersani et al., 2015).

Subsequently, one human study found that shorter leukocyte TL predicted worse antidepressant response to an SSRI (Hough et al., 2016). To the extent that accelerated biological aging is associated with antidepressant response, this effect does not seem likely to be confined to a specific treatment modality or drug. Shorter TL may also predict worse antidepressant response to pioglitazone, which also has anti-diabetic effects (Racagni et al., 2016). Turning to other disorders, in a group of bipolar and schizophrenia/psychosis patients, lithium non-responders had shorter telomeres than responders (Martinsson et al., 2013), and similar associations were found linking short TL to poor treatment response in schizophrenia (Li et al., 2015; Yu et al., 2008). No studies have investigated the association between TL and response to psychotherapy (e.g., cognitive behavioral therapy), although preliminary evidence suggests a positive correlation between mindfulness/meditation practices and telomere biology, including increased telomerase activity (Conklin et al., 2018; Schulte and Malouff, 2014). Furthermore, an increased inflammatory response may hamper the responsiveness to mood disorder treatments, and higher baseline inflammation may lead to treatment resistance (Rasgon et al., 2018). Moreover, an increased inflammatory response may also predict shorter telomeres (Martinsson et al., 2013). Changes in ccf-mtDNA levels were also found to be associated with SSRI treatment response, with the non-responders showing an increase in ccf-mtDNA and responders not changing (Lindqvist et al., 2018). A shared genetic disposition for inflammation, psychopathology, and treatment responsiveness has also been suggested (Zwicker et al., 2018). Furthermore, inflammatory indicators may provide information for potential personalized antidepressant recommendations, and could eventually guide the development of novel antidepressant treatments (Jha and Trivedi, 2018). Overall, the link between biological age and treatment responsiveness is a young field requiring further research.

5.3. Can biological aging be reversed with treatment?

One of the most relevant clinical questions in the field of biological aging and psychiatry is whether accelerated biological aging is a permanent imprint or a reversible process. While this may differ between indicators of biological aging, reversibility is at least possible to some extent for some indicators. For TL, the main mechanism of restoration is likely telomerase activation. Animal and in-vitro research provided evidence that telomerase-associated recovery of TL is to some extent possible (Batista et al., 2011; Jaskelioff et al., 2011). Several intervention studies have attempted to influence TL in humans (see Verhoef et al., 2014a, b for an overview). For example, recent small controlled studies have shown elongation of telomeres in response to highly controlled aerobic exercise verified with actigraphy (Puterman et al., 2018), losing and maintaining a weight loss of 10% or greater (Mason et al., 2018), and meditation-based interventions (Conklin et al., 2018). As mentioned above, for the field of psychiatry telomerase activation may be a mechanism of antidepressant action. Although no strong conclusions can be drawn due to a lack of well-powered clinical studies, there are several potential mechanisms by which psychiatric medications might modulate telomerase activity or TERT expression, including via increased brain-derived neurotrophic factor (BDNF) expression (Bersani et al., 2015). Increased telomerase activity may, in turn, induce clinical effects by promoting cell survival and/or functioning. It should, however, be noted that increased telomerase activity can render a cell immortal by preventing cell senescence, which under normal conditions prevents the transoformation of normal cells into cancer cells, and thus excessive telomerase activity is also associated with increased risk of cancer. Thus, while telomerase activation may hold the promise of reducing the risk of aging-related disease, the risk of less common but very serious cancer outcomes must be carefully weighed (Blackburn et al., 2015).

Cross-sectional studies show that certain behaviors may provide some protection against brain aging. Higher physical activity levels are associated with lower brain age (Steffener et al., 2016), and “younger brains” are seen in those that learn to play an instrument (Rogenmoser et al., 2018) and those who have practiced meditation for long periods (Luders et al., 2016). It remains to be elucidated whether brain age is responsive to intervention, but a randomized controlled trial (RCT) showed that ibuprofen temporarily reduced brain-PAD by 1.1 years in healthy individuals, likely due to its acute anti-inflammatory effects (Le et al., 2018). Physical activity may slow age-related DNA methylation changes in humans (Ren et al., 2012; Voisin et al., 2015). Longitudinal data shows that increasing BMI is associated with increasing epigenetic age (Quach et al., 2017), but in another report, epigenetic age from the liver was not “decelerated” after successful weight loss over a 9-month period through bariatric surgery (Horvath et al., 2014). Quach et al. (2017) also noted that consumption of fish, fruits, and vegetables, as well as effects of moderate alcohol, education, and income and exercise-induced anti-aging effects based on epigenetic age (Hannum clock). However, these findings are cross-sectional observations rather than longitudinal effects from RCTs. Nonetheless, a recent RCT suggests that vitamin D supplementation may decrease epigenetic aging based on the Horvath, but not Hannum epigenetic clock (Chen et al., 2018).

More controlled intervention studies are needed to determine whether biological aging indicators are truly modifiable in response to exercise, nutritional and/or pharmacological interventions. A common problem to observational studies is that behaviors tend to correlate, making it difficult to evaluate the specific influence of a given intervention or behavior in isolation. Individuals who exercise more tend to practice meditation more frequently, eat more plant-based and vegetarian diets, consume less illicit substances, etc. An additional problem with observational studies is that it is impossible to establish the direction of effects; while exercise may attenuate indicators of aging, biologically younger individuals may be more able and inclined to exercise.

6. Key challenges and priorities for future research

There are key challenges to accurately measure and interpret biological age indicators and to further our understanding of stress and biological aging. A partial list of six major challenges related to priority areas for the field, as well as recommendations to overcome them is presented in this section. We also summarize essential steps and propose a minimum standard for the design, collection, processing, analysis, and reporting of data involving biological age indicators (Fig. 4).

6.1. Challenge 1. Correlation is not causation

It is hazardous to infer causality from cross-sectional correlational data (Simons, 2015). For example, in the case of TL, it is possible that telomere shortening reflects states of stress, or responds to somatic or psychiatric illness, or at least to biochemical abnormalities associated with these states. It is also possible, however, that telomere shortening precedes somatic or psychiatric illness (Gotlib et al., 2014) or even underlies biological changes that cause these conditions. Biological age indicators could
be entirely independent of causal assessments of biological mechanisms that directly drive the aging process. Changes in biological age indicators could be the “canary in the coal mine” (Effros, 2009), representing factors associated with aging rather than aging itself.

Recommendation 1: Collect data longitudinally, consider experimentation, and choose prediction over explanation. This is a three-part recommendation. First, future studies should include longitudinal designs to increase the reliability and accuracy of measuring aging (Moffitt et al., 2017). Longitudinal measurements will also be important in determining which variables are critical for the maintenance of successful aging throughout the lifespan (e.g. absolute levels, change, variability), the “recipe” of which may vary between individuals (Sanders et al., 2012). Moreover, it will be critical to determine the optimal intervals of time between repeated assessments that are needed to detect meaningful changes in specific biological age indicators (see Fig. 3). This information about the timing and spacing of repeated measures, and real-life constraints should then be used to inform the design and choice of outcome measures when evaluating the effectiveness of interventions aiming to influence biological aging.

Second, experimental approaches utilizing cellular (or animal) models allow the direct manipulation of a specific (set of) variable(s). Thus, if we assume that a given stressor or predictor can be modeled accurately in vitro, experimental designs can provide direct causal evidence that a given factor is necessary and sufficient to produce a given outcome of interest. For this approach to empirically support a biological interaction between stressors and biological aging, the biological age indicator also needs to be detectable and meaningful in vitro, such as epigenetic age (e.g., Horvath et al., 2018). In cases where experimental demonstration is not possible, statistical methods such as causal inference (Bind et al., 2017) and genomic methods including Mendelian randomization (Burgess et al., 2012; Wiium-Andersen et al., 2017) can substantially reinforce our confidence regarding the direction of effects.

Third, in some cases, the number of predictors one wishes to consider is very large, either because there is no prior knowledge of their relative importance, or because the problem is truly complex – such as human biological aging. In such cases, the number of predictors can be large relative to the number of individuals, providing insufficient power for traditional inference-based statistics. In such cases, machine learning-based predictive modeling may be advisable to discover and validate predictive relationships between predictors and dependent variables. Whereas statistics draw population inferences from a sample, machine learning identifies generalizable predictive patterns (Bzdok et al., 2018). Statistical regression models traditionally used in psychological and epidemiological sciences are mathematically designed to maximize the explained variance (i.e., model fit) between sets of variables, without regard for generalizability to sub-samples or other samples. In statistical models, the stronger the fit the better the explanation. In contrast, predictive modeling approaches are designed to identify combinations of variables that together predict a particular health outcome, while simultaneously cross-validating combinations of predictors through either “in sample” or “out of sample” validation to prevent overfitting (overfitting decreases generalizability and increases the probability of spurious associations). Therefore, validated predictive models increase the likelihood, although they do not guarantee, that the identified predictors of biological age are robust, specific, and generalizable. An excellent article on the value of prediction over explanation in the psychological sciences is Yarkoni and Westfall (2017).
Regardless of the analytical approach taken, we should emphasize the value of converging evidence collected using different methods, measuring multiple (related and unrelated) predictors in parallel, and assessing multiple outcomes (Munafò and Davey Smith, 2018).

6.2. Challenge 2. Single biological age indicators are not correlated and may be better integrated

Not all measurements of biological aging are equally useful or interrelated, and it remains to be elucidated if and how different indicators relate to one another and which biological determinants are consistent across measures. As previously noted (Cole et al., 2018), no single biological age indicator can currently fully capture the complexity of the aging process, nor predict future health outcomes or lifespan with sufficient accuracy. There is therefore a need for combined indices that logically integrate multiple indicators (Fig. 5), hopefully resulting in accurate integrative panels that outperform single measures of biological aging, also previously suggested by Xia et al. (2017). Nevertheless, integration of indicators cannot compensate for inadequately powered studies, as large sample sizes are needed to ensure high generalizability.

Recommendation 2: Combine machine learning and other artificial intelligence techniques to create composite indices and panels of biological age indicators relevant to mental health. To date, several efforts to develop combined indices have been presented. For example, Marioni et al. (2016) showed an additive effect of combining TL and epigenetic age in explaining the proportion of age variance of their model. Similarly, Cole et al. (2017) explained significantly more variance in the prediction of mortality by combining brain-PAD and the Horvath epigenetic predicted age difference, than either indicator alone. Other examples include multivariate indicators of aging incorporating multiple physiological and functional measures (Belsky et al., 2015) and indices integrating multiple enzymatic and molecular measures of mitochondrial content and function in blood leukocytes as the MHI in association to psychological states (i.e., positive mood) (Picard et al., 2018). These constitute early attempts to reverse our reductionist inclinations and to move towards integrative metrics that will hopefully lead to improved prediction.

6.3. Challenge 3. Biological aging may be tissue- and cell-type specific

Related to the above, biological age indicators derived from specific cell or tissue types may not generalize to other cells or tissues. Biomarkers assessed in blood, for example, represent the average of multiple heterogeneous cell types. But cellular or organinal health may be more closely related to or reflected by indicators within individual cell types or those with the most extreme values (e.g., Flow-FISH, Aubert et al., 2012). Furthermore, TL differs across leukocyte cell types, such as naïve vs mature T cells (Chou and Effros, 2013; Lin et al., 2016), and differs in different regions of the brain (Mandani et al., 2015). Moreover, TL of particular cell types can be differentially vulnerable to attrition or affected by stress-related pathology (Boeck et al., 2018b).

Recommendation 3: Purify cell types using established molecular markers. Purification of cell types can be accomplished by a variety of methods (flow cytometry, magnetic-activated cell sorting), yielding living cells amenable to downstream molecular and cellular analyses (Lin et al., 2010). Under certain conditions where it is not possible to isolate specific cell subtypes, it may be difficult to interpret certain indicators that exhibit large cell type-specific differences (such as mtDNAcn in blood). In some limited cases where a lot of information is available for adjustment, such as for DNAm measured on bead chips (100,000’s of data points), it is possible to use statistical approaches, i.e. reference-based (Hattab et al., 2017) or reference-free (Housman et al., 2014), to infer underlying cell type proportions (Titus et al., 2017) and adjust results accordingly. Another interesting application is demonstrated in a preliminary study by Chan et al. (2018) that uses deconvolution approaches to show novel MDD-methylation associations in individual sub-populations of neurons/glia from bulk brain, as well as in granulocytes/T-cells/B-cells/monocytes from bulk blood data. In addition, creative ways to harvest other cell types from tissues other than blood could yield increasingly meaningful biological age indicators. For example, the DNAm-based skin & blood clock by (Horvath et al., 2018) is robust across tissues (e.g. fibroblasts, buccal, endothelial, saliva samples) and can, therefore, be applied to many organs, as well as ex vivo. Thus, this approach provides extended information on synchronized biological aging, independent of sampling source.

6.4. Challenge 4. Within-group variance may be larger than between-group variance

Although group differences in biological age indicators have been reported in several psychiatric illnesses (Darrow et al., 2016; Lindqvist et al., 2015), these reflect average group differences. However, there is often considerable within-groups variability and considerable overlap between groups, making it very difficult to use most biological age indicators as diagnostic aids. In addition, even specific psychopathological diagnoses (e.g., MDD or schizophrenia) often include individuals that vary widely in their symptoms and presentation, making the search for predictors of mental illnesses defined diagnostically somewhat elusive (Cutbert and Insel, 2013). As Wallers et al. (2018) suggest, the complex, highly polygenic and multifaceted causes of severe mental disorders may only be fully understood by mapping patients’ individual signatures, rather than studying the average patient. Population-based normal ranges have yet to be reliably determined, so biological age indicators may be more useful in detecting within-subject changes over time rather than in comparing individuals or in establishing actuarial norms. Nonetheless, a recent study utilizing flow-FISH assay techniques

Fig. 5. Topology of biological age indicators, upstream measures, and downstream integrative indices and panels. Biologically-informed and functionally relevant composite indices integrating two or more indicators can be derived from individual indicators. Their added predictive power should be validated by either in-sample cross-validation or preferably out-of-sample validation. Integrative biological age panels may eventually outperform single biological age indicators and indices due to the strength of their association, their construct stability, and/or their greater generalizability across individuals and independent samples.
reported reproducible and definable upper and lower normal boundaries for TL in a hospital population (Alder et al., 2018), indicating that standardization of these measures may be achievable. Within-group variance is often due to measurable individual differences in behavior, health condition, and lifestyle.

**Recommendation 4:** Visualize your data, carefully assess known influences, move beyond group-based analyses, and use within-person modeling approaches. To visualize data where there are multiple measures over time per individual, we advocate for spaghetti plots (example with epigenetic age and TL can be found in Marioni et al. (2016), and with cortisol trajectories in Ram and Grimm (2009). Standard measures of variance can also be used to model population variability at different timepoints and to compare sub-groups, and more sophisticated mathematical approaches can be useful to generate individualized phenotypes (Hertel et al., 2018). Data reduction approaches such as principal component analysis (PCA), partial least square discriminant analysis (PLS-DA), and t-distributed stochastic neighbor embedding (t-SNE) are also useful to visualize high-dimensional data in two or three dimensions and to assess whether subsets of individuals in the sample naturally cluster together (van der Maaten and Hinton, 2008; Xia et al., 2015). This kind of approach can provide evidence of shared phenotypes or trajectories that would otherwise remain undetectable by standard uni- or multivariate analyses. Other statistical approaches to identify different subgroups exhibiting different trajectories in biological aging or in clinical course include growth mixture modeling (Ram and Grimm, 2009), and random coefficients linear regression models to examine within-person changes over time (Diaz et al., 2018).

**6.5. Challenge 5. Large sample sizes are needed to detect small effect sizes**

Related to the point above, group differences between individuals with certain psychiatric illnesses vs. healthy individuals, even if statistically significant, often have small effect sizes that require large sample sizes to be demonstrated (Darrow et al., 2016; Han et al., 2018; Verhoeven et al., 2014a, b; Whalley et al., 2017). Similarly, even using predictive modeling and machine learning approaches, small sample sizes are more susceptible to overfitting (Yarkoni and Westfall, 2017). It would be informative to pool databases across studies and to check for consistency and predictive accuracy across studies, albeit at the cost of increasing heterogeneity of the samples studied and of the laboratory methods used. (Schnack and Kahn, 2016) argue that larger sample sizes will have more generalization power. This is important for creating robust “canonical” publicly available prediction models (e.g. Horvath’s epigenetic clock) that can be readily applied to smaller studies that cannot permit partitioning their data into a training and validation set.

**Recommendation 5:** Collaborate and harmonize data collection and analysis protocols to facilitate data pooling worldwide. Certain types of data may be available through crowdsourcing, a data collection process with remarkable scalability that may enable identification of robust small effect size effects (Mohammadi, 2015). Large-scale collaborative initiatives like the ENIGMA consortium for imaging and genetics data may be useful examples (Thompson et al., 2014). When not possible to replicate findings on sufficiently large independent datasets, use cross-validation methods to avoid overfitting (Yarkoni and Westfall, 2017).

**6.6. Challenge 6. Non-uniform laboratory assays and storage conditions**

Different assay methodologies can yield relatively different results (e.g. qPCR vs. Southern blot for TL assessment, Aviv et al., 2011), and most assays such as Flow-FISH assays and qPCR ascertain relative TL rather than absolute length. These issues have been discussed in detail elsewhere (Aubert et al., 2012), and the relative merits of the different assays have been compared. A more mundane but important and under-appreciated caveat is that methodological differences can yield spurious results (e.g., length of time a specimen was kept in a freezer, freezer temperature, whether whole blood or intact cells vs. cellular lysate were frozen, specific batch of reagents used, assay technique, method of DNA extraction [for TL] (Dagnall et al., 2017; Khincha et al., 2017; Shiels, 2010). An important issue that has not yet been experimentally assessed in multi-year human studies is the relative merit of assaying batches of samples at short intervals (e.g., every year) to minimize freezer time, compared to keeping all samples frozen until the end of the entire multi-year study so that all samples can be assayed simultaneously using identical procedures and reagents.

**Recommendation 6:** Harmonize measurements and storage conditions - colder is better. Follow guidelines, where available, to ensure that samples are collected and measured with the highest standards. Efforts are currently underway to systematically compare methods available to measure TL and will hopefully produce specific guidelines that can be implemented at a large scale. Fig. 4 summarizes some measures that can be considered in the design, collection, pre-processing, analysis, and reporting of data to give researchers the ability to critically evaluate published results and hopefully harmonize methods and datasets.

7. **Summary**

Much progress has been made in the last decade towards developing objective biological age indicators but several key challenges and opportunities remain. Multiple new indicators spanning physiological and functional capacity, brain function and structure, and cellular and molecular levels of analysis have recently emerged, particularly under the force of ‘omics’ technologies. Most indicators show good to excellent correlations with chronological age, and some have the ability to predict age-related outcomes such as mortality with moderate accuracy (see Table 1). However, some have not yet been prospectively studied in large cohorts so their predictive power in relation to health outcomes remains unknown. In fact, many proposed biological age indicators still require replication and validation in larger independent datasets. As a whole, the development of multi-systemic biological age indicators has demonstrated that aging occurs not at a single level in a specific cell type, but rather manifests somewhat differently in various organs and tissues, cell types, and across a number of levels (see Fig. 1). Rapidly developing methods show that biological age indicators gain in precision and prediction accuracy by leveraging biologically-informed approaches to integrate individual indicators into more powerful indices and panels of measures.

In relation to psychological stress and mental health, much work also remains to establish to what extent and how psychological states influence the aging process. Equally important to mapping these psycho-biological processes is to understand the “reverse” causal link whereby accelerated biological aging may impact physiological vulnerability and resilience to life stressors and psychopathology. A shared objective for our field is to disentangle these associations and identify the causal pathways that drive aging and mental health trajectories. To do so, we need to address number of key challenges that, if adequately met, will yield exciting opportunities to advance our understanding of human health.

To achieve this objective, it will also be essential for clinicians, psychologists, epidemiologists, and behavioral scientists to enter in a dialogue with experimental biologists. Such dialogue immediately opens new questions that would not otherwise arise within the silos of individual disciplines, departments, and laboratories. For example, the joining of ideas around psychosocial stress and cellular aging (Epel et al., 2004), age-related elevation in circulating DNA (Teo et al., 2018), and of psychoneuroimmunology and the pro-inflammatory effects of psychological stress (Marsland et al., 2017) has led to the idea that acute psychological stress may rapidly trigger immunogenic mtDNA release into the circulation (Trumpf et al., 2019). If met with sufficient enthusiasm and resources, these jointly-created interdisciplinary questions can subsequently push the development of new laboratory methods, statistical and analytical tools, and new theoretical models. The interdisciplinary field of human psychobiology is replete with opportunities for innovation and discoveries.
Overall, we believe that developing, validating, and studying predictive biological age indicators – and defining how psychological states or psychiatric illnesses influence them – will fill important knowledge gaps linking stress and mental illness to aging. Achieving this goal will have at least three main positive consequences for the biomedical sciences generally. First, it will enable the stratification of individuals (and group of individuals) based on their current health state and future disease risk, including symptoms, disorders. This notion converges with the precision medicine agenda that aims to identify individualized predictors of future mental health and disease states (Insel and Cuthbert, 2015; Abi-Dargham and Horga, 2016). Achieving this goal could benefit clinical practice by providing clear and objective guidelines to direct health-promoting interventions and treatment delivery in a personalized way, specifically one that is most aligned with the individual’s needs (Picard et al., 2013). Second, achieving this goal will provide objective and sensitive methods to evaluate the effectiveness of health-promoting interventions. The efficacy of interventions aimed at decreasing or even reversing advanced or premature aging, as well as interventions aimed at enhancing well-being and other health outcomes would be more efficiently assessed using precise and sensitive biological age indicators. Finally, mapping stress-sensitive biological age indicators will also generate knowledge about human aging that can be taken back to the laboratory bench to orient basic biological science. Specifically, understanding the basis of normal and abnormal aging processes may identify new physiological targets for therapeutic intervention. Thus, an unintended consequence of improving existing biological age indicators and developing new end points for clinical and epidemiological human research may be, in turn, create leaps in understanding about the fundamental biological mechanisms that explain why we age in the first place.

Conflict of interest

The authors have no conflict of interest to report.

Acknowledgements

This paper was inspired by a focus group meeting hosted by J.E. Verhoeven and B.W.J.H. Pennuth at Amsterdam UMC, Vrije Universiteit, Department of Psychiatry in Amsterdam in June 2018. The authors are grateful to investigators and patients whose work has enriched this review, and to the anonymous reviewers for their constructive criticisms. Work of the authors is supported by the Wharton Fund, NIH grants GM119793 and MH113011 (MP), HD086487 and MH101107 (ART), MH083784 (OMW).

References

Akulenko, R., Merl, M., Helms, V., 2016. BEclear:batch effectdetection and adjustment in...
bioRxiv.


Major Depressive Disorder: Results from the ENIGMA Major Depressive Disorder working group. bioRxiv.


Franke, C., Gaer, C., 2012. Longitudinal Changes in Individual BrainAGE in Healthy Aging. Mild Cognitive Impairment, and Alzheimer’s Disease 1Data used in preparation of this article were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). GeroPsych 25, 225–245.


Han, L.K.M., Dinking, R., Hahn, T., Ching, C., Eyler, L., Aftanas, L., 2019. Brain Aging in Major Depressive Disorder: Results from the ENIGMA Major Depressive Disorder group. bioRxiv.


